REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
5.Sep.02

3. REPORT TYPE AND DATES COVERED
THESIS

4. TITLE AND SUBTITLE
EXPERIMENTAL STUDY OF AUTOMATION TO SUPPORT TIME-CRITICAL REPLANNING DECISIONS

5. FUNDING NUMBERS

6. AUTHOR(S)
2D LT JOHNSON KIP E

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

8. PERFORMING ORGANIZATION REPORT NUMBER
CI02-518

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

20021015 107

14. SUBJECT TERMS

15. NUMBER OF PAGES
133

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT

18. SECURITY CLASSIFICATION OF THIS PAGE

19. SECURITY CLASSIFICATION OF ABSTRACT

20. LIMITATION OF ABSTRACT

EXPERIMENTAL STUDY OF AUTOMATION TO SUPPORT TIME-CRITICAL REPLANNING DECISIONS

By

KIP EDWARD JOHNSON

Submitted to the Department of Aeronautics and Astronautics
on May 17, 2002 in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Aeronautics and Astronautics

ABSTRACT

An experimental study was performed on degrees of automation to support time-critical decision-making in complex environments, specifically the in-flight replanning task. Fourteen subjects interacted with a part-task military combat simulation of the in-flight replanning task. Subjects modified a two-dimensional route through waypoint manipulations on a computer monitor, in response to a sudden change in the simulated flight environment.

The study focused on determining the relationships between automation assistance, time pressures, and information elements as inputs, and the resulting decision performance as outputs. In addition to a baseline case without automation (None), subjects received one of three types of route-assistance automation, which provided a static route suggestion in response to the environmental change. The route-assistance automation either reduced hazard exposure (Hazard), ensured the satisfaction of fuel and time-on-target constraints (Constraint), or combined the two (Full). The experiment exposed subjects to four time-pressured conditions—20, 28, 40, and 55 seconds—and to a condition without time pressure.

Using a route cost metric, overall route cost with Full (0.123 average) and Hazard (0.406 average) automation assistance was significantly lower (better) than without automation (0.468 average); this demonstrated that automation assisted subjects in the replanning task. Most benefit from automation occurred in the highly time-pressured conditions; at lower time pressures, performance with None was similar to conditions with automation assistance. The benefit from Full was nearly double the sum of its individual Hazard and Constraint module benefits, with a benefit difference of 0.212 at each time pressure. There was a 14.3 percent mission failure rate, with 32 failures out of 224 total trials. There were more mission failures with Hazard (13 failures) and with Constraint (8 failures), than with None (6 failures); while the least failures occurred with Full (5 failures). This showed that automation of partially integrated information could induce certain problems not otherwise observed in cases without automation. Lastly, without time pressure, subjects outperformed any time-pressured trial, even with automation assistance.

Thesis Supervisor: James K. Kuchar
Title: Associate Professor of Aeronautics and Astronautics

Experimental Study of Automation to Support Time-Critical Replanning Decisions

by

KIP E. JOHNSON

B.S. Aeronautical Engineering
U.S. Air Force Academy, 2000

Submitted to the Department of Aeronautics and Astronautics
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Aeronautics and Astronautics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2002

Signature of Author: ______________________________
Department of Aeronautics and Astronautics
May 17, 2002

Certified by: ______________________________
James K. Kuchar
Associate Professor of Aeronautics and Astronautics
Thesis Supervisor

Accepted by: ______________________________
Wallace E. Vander Velde
Professor of Aeronautics and Astronautics
Chair, Committee on Graduate Students

[Page Intentionally Blank]

EXPERIMENTAL STUDY OF AUTOMATION TO SUPPORT TIME-CRITICAL REPLANNING DECISIONS

By

KIP EDWARD JOHNSON

Submitted to the Department of Aeronautics and Astronautics
on May 17, 2002 in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Aeronautics and Astronautics

ABSTRACT

An experimental study was performed on degrees of automation to support time-critical decision-making in complex environments, specifically the in-flight replanning task. Fourteen subjects interacted with a part-task military combat simulation of the in-flight replanning task. Subjects modified a two-dimensional route through waypoint manipulations on a computer monitor, in response to a sudden change in the simulated flight environment.

The study focused on determining the relationships between automation assistance, time pressures, and information elements as inputs, and the resulting decision performance as outputs. In addition to a baseline case without automation (None), subjects received one of three types of route-assistance automation, which provided a static route suggestion in response to the environmental change. The route-assistance automation either reduced hazard exposure (Hazard), ensured the satisfaction of fuel and time-on-target constraints (Constraint), or combined the two (Full). The experiment exposed subjects to four time-pressured conditions—20, 28, 40, and 55 seconds—and to a condition without time pressure.

Using a route cost metric, overall route cost with Full (0.123 average) and Hazard (0.406 average) automation assistance was significantly lower (better) than without automation (0.468 average); this demonstrated that automation assisted subjects in the replanning task. Most benefit from automation occurred in the highly time-pressured conditions; at lower time pressures, performance with None was similar to conditions with automation assistance. The benefit from Full was nearly double the sum of its individual Hazard and Constraint module benefits, with a benefit difference of 0.212 at each time pressure. There was a 14.3 percent mission failure rate, with 32 failures out of 224 total trials. There were more mission failures with Hazard (13 failures) and with Constraint (8 failures), than with None (6 failures); while the least failures occurred with Full (5 failures). This showed that automation of partially integrated information could induce certain problems not otherwise observed in cases without automation. Lastly, without time pressure, subjects outperformed any time-pressured trial, even with automation assistance.

Thesis Supervisor: James K. Kuchar
Title: Associate Professor of Aeronautics and Astronautics

[Page Intentionally Blank]

Note:

This thesis was prepared at the International Center for Air Transportation, Massachusetts Institute of Technology (MIT), under the Office of Naval Research Award # N00014-00-1-0659. The period of research spanned from January 2001 to February 2002. The views expressed in this document are those of the author, and do not constitute the official policy or position of the U.S. Air Force, Department of Defense, U.S. Government, Office of Naval Research, or MIT.

[Page Intentionally Blank]

Table of Contents

List of Figures

List of Tables

List of Acronyms, Symbols, and Abbreviations

Acronyms

TCAS .... Traffic Collision and Avoidance System
DM .... Decision-Making
NDM .... Naturalistic Decision-Making
PA .... Pilot's Associate
RPA .... Rotorcraft Pilot's Associate
CASSY.................Cockpit Assistance System
RTIP ....Real-Time Integrated Planner
DIR .... Dynamic In-Flight Replanner
TOT ....Time-on-Target
GLUT .... OpenGL Utility Toolkit
SE .... Software Editor
ANOVA .... Analysis of Variance
MS .... Mean Square
SD .... Standard Deviation
CT .... Characteristic Time

Symbols

nm .... Nautical Miles
ln .... Natural Log
p .... Probability Value
F(#,#).................. F Statistic (numerator degrees of freedom, denominator degrees of freedom)
z .... Standard Normal Statistic
χ^2 .... Chi-Square

Abbreviations

None .... No Automation Assistance
Partial .... Partial Automation Assistance
Full .... Full Automation Assistance
Constraint ....Constraint Automation Assistance
Hazard .... Hazard Automation Assistance
rot .... Rotation
sec .... Seconds
min .... Minutes
Time .... Time Pressure Information Element
Hazard..................Hazard Information Element
Fuel .... Fuel Information Element
TOT .... Time-on-Target Information Element

1. INTRODUCTION

Complex, uncertain, and time-critical environments filled with a plethora of diverse and dynamic information elements continually push the limits of human sensory and cognitive ability, driving the need for automated decision-support systems. While there is a clear need for automation that reduces human cognitive workload, designing an effective automated decision-aiding system is a difficult task [Parasuraman & Riley, 1997; Sexton 1988; Sarter & Woods, 1995]. Unstructured and uncertain aspects to a problem, with multiple competing interests and goals, characterize these complex environments. Full and complete automation may not be appropriate or feasible for complex environments because automation often does not have access to or cannot accurately model relationships between all relevant information [Parasuraman & Riley, 1997; Scerbo, 1996]; rather, automation working in parallel with a human for decision-making tasks would be more appropriate [Taylor & Reising, 1998; Schulte, et al., 1999; Layton, et al., 1994; Aust, 1996]. Decision-support systems should take advantage of the human's ability to make value and risk judgments in the face of competing factors that may constrain a problem's analytical solution. Therefore, some form of cooperation between human and automation is generally required.

Under extreme time pressures, well-designed automation assistance should provide at least some benefit to the human because there is no time for the human to form productive decisions. However, as this time pressure relaxes, the benefits from automation assistance may decrease, possibly even to the point of hindering human decision-making performance. This automation hindrance would partly be due to the need to first understand and decompose the automated solution. In the environment where automation cannot observe everything, the unaided human will hypothetically perform as well as or better than with automated assistance given enough time. Understandably, significant automation design issues exist as to how much and what type of information to process when suggesting a solution, and to what degree the human should be involved in the decision-making process.

Figure 1-1 is a general model illustrating an automated decision-support task, where the human is part of the decision-making and actuation process. Let us step through this model. First, sensors filter information from the environment for both the human and automation to use.

Humans usually receive information from the sensors through visual or audio displays. Next, the human and automation both process and analyze the sensed information, using some level of decision-making collaboration. Collaboration may range from complete automation assistance to no automation assistance. In addition, collaboration may be fixed or adaptive to time pressure, problem complexity, or human desires for example. The appropriate design of automated decision-support is imperative for an effective collaboration with humans. Finally, the automation or human makes a decision and executes an action.

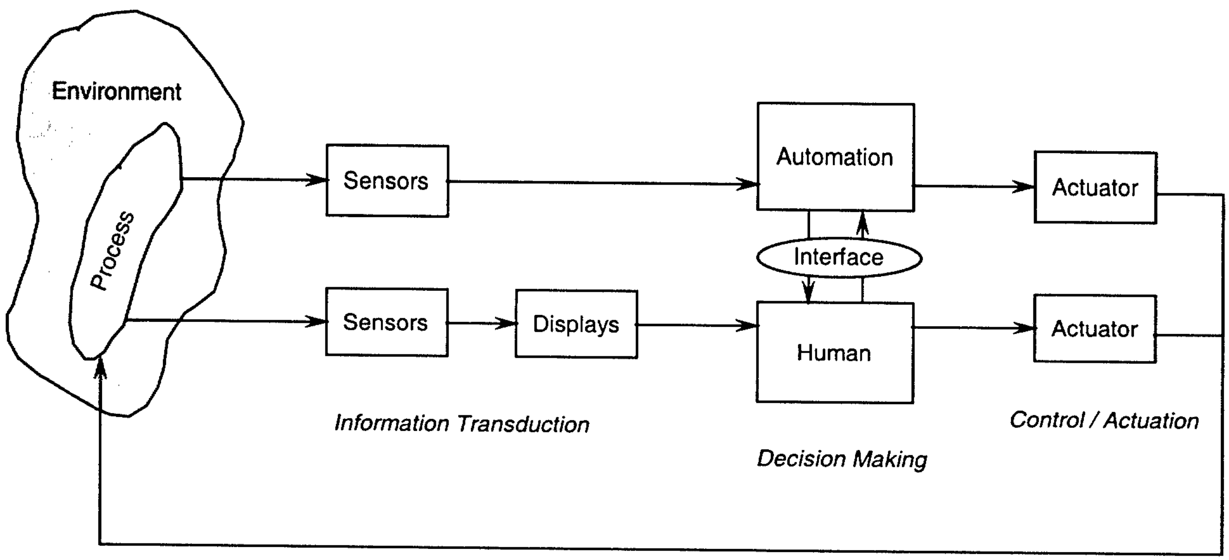


Figure 1-1. Decision-Support Task Model (courtesy of J.K. Kuchar).

The time available to the human for decision-making can range from immediate (few seconds), to tactical (few minutes), to strategic (greater than a few minutes), depending on the environment's process demanding the solution [Fan, et al., 1998]. For our research, the complex and dynamic environment was an air-to-ground combat-flight mission, with an in-flight replanning task as the process of interest. For example, a dynamic thunderstorm may demand a route replanning solution on the order of a few minutes (tactical), while an anti-aircraft missile launch most likely demands a replanning solution within a few seconds (immediate).

Formal and quantitative goals, with associated strategies, drive decision-making behavior for automation, while this is not always the case for humans. What information goes into the decision-making process for in-flight replanning? Following is a list of information elements considered for both commercial and military aviation scenarios, but this invariably is not all-

inclusive. The information elements were broken into strategic, tactical, and immediate temporal decision-making categories, and could be defined in more than one category. These information elements represented constraints, hazards, and goals, of which some are quantifiable, and many others are ill structured and qualitative.

Strategic:

- Weather hazards [Latorella & Jenkins, 1999]—turbulence, convection, icing, volcanic activity, ozone concentration
- Winds aloft
- Fuel efficiency
- Passenger comfort
- Arrival time
- Traffic congestion
- Runway/Airport closure
- Restricted airspaces
- Destinations/targets

Tactical:

- Thunderstorms (convection)
- Traffic congestion
- Emergency—passenger, hydraulics, other aircraft emergencies, icing
- Low level wind shear
- Wind gusts

Immediate:

- Traffic hazard
- Terrain hazard
- Emergency—fire, engines out, icing, cabin depressurization

Military Specific:

Tactical:

- Destination change (due to a change in target hierarchy, the reconnaissance area, or threats)
- Fuel level

Immediate:

- Threat exposure
- Threat movement
- Target movement
- Target kill
- Fuel level

Both humans and automation would use the above information elements when making decisions regarding flight management, which ultimately result in defining the aircraft route trajectory, whether pre-planning on the ground or in-flight replanning. Executing an in-flight replan would involve selecting the aircraft's current state, such as heading, velocity, and altitude, and planning for future states. Invariably, decision-making for the in-flight replanning task will be a collaborative effort between humans and cockpit automation.

How should we design automation to capitalize on an individual's decision-making abilities? This document addresses an experimental study performed to determine the most effective design of decision-aiding automation for time-critical tasks, and within complex and uncertain environments.

1.1 Problem Statement

Three properties define the problem category of interest to our research. First, the decision-aiding automation does not have access to all the information available to the human. Conceptually it would be difficult to model all information a human could process into an automated decision expert system, let alone be possible to implement complete automation of cognitive tasks in actual practice [Parasuraman & Riley, 1997]. For example, a traffic collision and avoidance system (TCAS) may alert the pilot to descend and avoid an approaching aircraft. However, the pilot could have supplemental information that another aircraft is below, in which descending as TCAS directed may induce a different collision. Therefore, the pilot chooses to

make a constant altitude right banking turn, avoiding both the approaching aircraft and the aircraft below. When the automation cannot see everything, due to sensor limitations for example, there is the possibility for a human to outperform the automation in the performance parameter of interest.

Second, the problem is within the time-critical domain for decision-making, defined as a time between the immediate and tactical decision-making time scales described above. The time-critical domain has the potential for many beneficial human and automaton interactions. With little time pressure (i.e., such as at the strategic time scales), the human should at least be able to produce a solution as good as the automated solution. When the process requires a decision in the sub-second to few seconds time scale, the human simply does not have enough time to make clear observations, let alone coherent decisions. Automation is valuable in these situations due to its ability to obtain and process information more quickly than the human. Within the time-critical domain, however, there is time for humans to make decisions, yet not enough time to make clearly informed and beneficial decisions without some automation assistance. The human interaction with automation can range from observation and monitoring of automated decisions to some degree of decision-making collaboration with the automation.

Lastly, the problem involves a decision-making task from a naturalistic environment, defined as an uncertain, complex, and dynamic environment, with multiple competing interests and goals [Cannon-Bowers, et al., 1996]. These problems make the design of automation for cognitive tasks, such as decision-making, a difficult endeavor. There is often too much uncertainty in the environment, and biases among individuals, that make it difficult to clearly design the automation through formal logic and rules. A naturalistic environment best highlights the human's innate ability to process information, without rules and restrictions, in the face of conflicting interests and goals to arrive at a solution. The study of decision-making using simple or easy environments is trivial, and would not have many practical applications.

This field of automation applications for time-critical decision-aiding is rich with the potential to benefit immensely from human factors research [Pritchard, 2000; Cannon-Bowers, et al., 1996; Boeing, 2002]. The research would need to focus on determining the most effective human and automation interactions for producing nearly optimal solutions under time constraints. We hypothesized that a replanner's automation should be intelligent, being able to filter and integrate the appropriate types and amounts of information based on the task

complexity, existing time pressure, and user performance. Too little automation may not adequately aid the human in developing a solution in time-constrained problems, but too much automation could reduce or even hinder the benefits from a human's intuition and ability to integrate diverse information without rule or logic-based constraints.

While there is plenty of work dedicated to automation and decision-making in relation to human performance or situational awareness, research combining the two specifically within a time-critical domain is limited. The one academic paper found relating automation to time-critical tasks focused more on automation of fault management in thermal-hydraulic systems [Moray, et al., 2000]. Also limited is the literature on intelligent, or "adaptive," automation in relation to decision-making. A recently completed study looked at the human performance differences between varying cognitive tasks, which included sensing, analyzing, decision-making, and action execution, when using adaptive automation [Kaber, et al., 2002]. This dearth of research is even more striking when realizing that adaptive automation for time-critical decision-aiding has countless applications in today's world: commercial and general aviation, military combat aviation, command and control of wartime operations, medical care, finances, and the control of energy and chemical production processes to name a few.

We conducted an experimental study of automation to support time-critical in-flight replanning decisions in complex environments. In-flight replanning in response to information updates can be a cognitively demanding task for a pilot, especially under time pressures and when consequences of poor decisions may result in the loss of lives. In efforts to better design pilot decision-support systems, which include assisting humans in the in-flight replanning task, this study was designed to answer the following questions:

1. Can we quantitatively measure, with accuracy, human replanning performance within a complex and time-constrained environment through a multi-variable human factors experiment?
2. How do interactions between time pressures and automation assistance types affect human replanning performance?
3. How do subjects perceive the automation assistance and their replanning performance?
4. How do the various information elements interact and affect the replanning process?

The experimental goals were twofold. Figure 1-2 shows the primary experimental objective: to objectively and subjectively determine the relationships between varying degrees and types of automation assistance, and time pressures as inputs; and the resulting human decision performance for the in-flight replanning task as outputs. We modeled individual human ability, problem complexity, time pressure, and route automation assistance as the primary influences on replanning decision performance. A route cost quality metric objectively measured human decision-making performance, while subjects provided the subjective results directly. In addition, we wanted to develop (or further refine) a generalized model for decision-support systems by combining automation and time pressure relationship findings with a better understanding of the important information elements present in the flight environment.

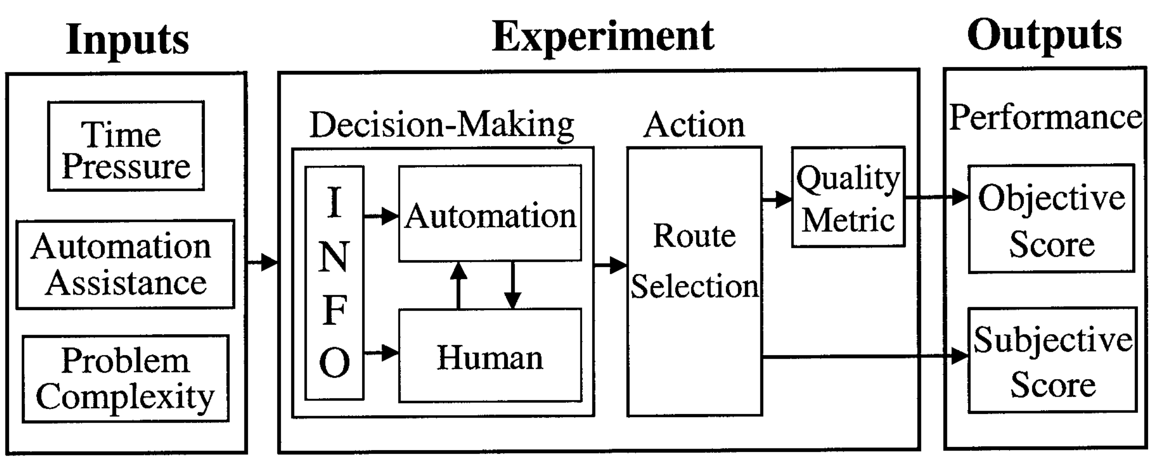


Figure 1-2. Experimental Overview.

Let us outline the rest of this document's contents. Chapter 2 gives a background on decision-making, humans and automation, and in-flight replanner technology, providing the foundation and motivation for our research. Chapter 3 provides a detailed description of the experiment design. Chapter 3 discusses the combat environment, the display interface, the experimental protocol, the experimental variables, the training, and the software. Chapters 4 and 5 exhaustively cover the quantitative and qualitative experimental results. Chapter 6, Discussion, combines the quantitative and qualitative results into ideas and findings that are more cohesive and not otherwise observed directly from the results. Finally, Conclusions Chapter 7 highlights the key experimental findings, discusses the applications of this research, and provides a direction for future research.

[Page Intentionally Blank]

2. BACKGROUND

Chapter 2 expands the discussion of important topics found in our research: time-critical decision-making in complex environments, the human and automation interaction, adaptive automation, and in-flight replanner technology. Theoretical and empirical research from numerous authors will be discussed to further explain and validate the importance of our experimental goals, aforementioned in the Introduction chapter. When applicable, the chapter highlights the shortcomings and potential areas that could benefit from research of automated decision-support for time-constrained, uncertain, and complex environments.

2.1 Naturalistic Decision-Making

Decision-making (DM) is a cognitive task defined as the processing of information with regard to choices, and the selection of one choice amongst the alternatives in the face of ambiguity, which requires a relatively long time scale (longer than a second) [Wickens, et al., 1998]. The effectiveness of automated decision-support systems relies on the accurate understanding and modeling of human decision-making. The lack of accuracy in understanding and modeling human behavior is arguably the reason for so many human-induced accidents when interacting with automation.

Modeling what a rational human "should" do forms the foundation of classical decision-making theory. As cited in Wickens et al. [1998], understanding classical decision theory is important because it is the foundation for many computer-based decision-aids. However, classical DM theory is limited in scope and does not model how humans actually behave. Often, humans do not behave as expected in complex and uncertain applications, in which probability theories and formal rules have defined what is normal behavior [Cannon-Bowers, et al., 1996]. In addition, classical DM theory does not account for the natural differences between humans in psychological processes and strategies. In light of these shortcomings, recent DM research focuses on developing descriptive models, based on heuristics, to describe actual human DM, called naturalistic decision-making (NDM).

Current NDM research strives to model human DM in real-world environments [Cannon-Bowers, et al., 1996]. These naturalistic environments are described by dynamic and uncertain tasks, complex and ill-structured problems, competing goals, and high stakes. Rasmussmen's skill-rule-knowledge based model is one NDM model that takes into account the nature of the task and the human's associated experiences with the task. This model distinguishes three levels for decision-making—skill, rule, and knowledge-based—in which a human can operate in each level depending on personal experiences and the novelty of the problem. Cannon-Bowers et al. [1996] recognized the lack of theory and empirical results in the NDM field, and urged researchers to fill this gap because of its importance to the design of decision-support systems.

The combat flight environment, for example, has too much uncertainty and complexity for any decision-aid to be designed by what a human "should" do. On the contrary, our NDM research focused on how humans actually behave in a generic and abstract replanning task. From this research, we wanted to develop a descriptive model of pilot behavior with respect to in-flight replanning, combining interactions with time pressures, automation assistance categories, information elements, and personal background experiences. We are trying to link human psychological processes and strategies to real world tasks, a connection that is necessary for the success of decision-support systems.

2.2 Humans and Automation

Automation is applied everywhere in today's world, from the automation of opening garage doors to the automation of error checking shuttle launch procedures. Automation can supplant physical activities. The motivation for such automation could be financial, the relief of labor-intensive or boring work, or safety for example. Automation can allow human interactions in physical control processes otherwise not possible, such as the stability augmentation systems that allow pilots to fly inherently unstable aircraft. In addition, automation can aid the human in performing cognitive functions, which include decision-making, monitoring, planning, or idea generating for example. The motivation for cognitive assistance automation is to increase human performance by reducing the overall cognitive workload and increasing situational awareness.

Our research focused on automation for cognitive functions, a focus that invokes much discussion and debate from a theoretical and empirical perspective. The definition of automation

is not clear, and it changes with time and technological innovations. For the purposes of this document, we define automation to be the partial to full execution of a system function by a machine that was or could have been carried out by a human [Parasuraman & Riley, 1997]. The computer is the primary machine agent for many automation applications today. The recognized challenge for system designers is what and how much to automate, and this answer is not always clear [Parasuraman, et al., 2000].

There are many examples of properly designed and integrated automated systems in all fields. In the aviation field for example, cockpit predictor displays have reduced workload and improved hazard detection performance, and the horizontal situation indicator has dramatically reduced pilot workload (as cited in [Parasuraman, et al., 2000]). The study and discussion of properly designed automation systems is of little concern here, however, because there is not an overwhelming impetus for change.

Both users and designers debate the benefits from automation in support of cognitive tasks, however. There are countless examples of poorly designed automation that negatively affected human cognitive performance and better illustrate the controversies in automation design. A number of controlled flight into terrain accidents demonstrated the need for continuing improvement of feedback on the current mode of automation in civil transport aircraft [Parasuraman & Riley, 1997]. A study by Sarter and Woods [1995] also showed that most of the human and automation interaction difficulties stem from a lack of mode awareness by the users (pilots). Parasuraman and Riley [1997] cited several examples where an under or overreliance on automation caused railroad and aviation accidents. A survey conducted by Wiener [1988] found conflicting responses from Boeing 757 pilots on whether cockpit automation reduced or increased total workload. A navigation expert system study demonstrated that out-of-the-loop performance was able to cause a loss in situational awareness and the degradation of manual skills as a direct result of information processing automation [Endsley & Kiris, 1995]. The aforementioned give a small glimpse at the main problems associated with the introduction of automation: complacency, skill degradation, increases in cognitive workload, and the loss of situational awareness.

Unfortunately, examples of poorly designed automation in the real world can have catastrophic results. The failure of automation in a nuclear plant or in the flight control system of an inherently unstable aircraft, for example, can result in the loss of lives. It is not correct,

however, to place complete blame on human error alone; rather, blame should be placed on human error due to inappropriate interactions with automation. These implications of automation design clearly show the necessity for human factors considerations in the design process of automated systems; however, where do we start?

With today's computer technological advances, there is little left that cannot be automated. The problem now is how to design the human-automation interaction to produce a clearly effective and beneficial system, a system that reduces total workload or increases situational awareness, without causing an unnecessary amount of skill degradation or complacency issues. To facilitate the design of effective automation, many researchers have proposed cognitive task models to answer this question: What function(s) of the cognitive task can and should be automated? These models strive to accurately describe and categorize the human-automation interactions of various tasks, ranging from general system level perspectives to more detailed models, such as the replanning task [Parasuraman, et al., 2000; Fan, et al., 1998].

With an understanding of the cognitive tasks, the automation designer now has another important consideration: to what level should each cognitive function be automated? Unfortunately, full automation of cognitive functions may not be the simple answer, and this is regardless of the fact that achieving effective full cognitive automation is a very difficult task for many applications [Parasuraman & Riley, 1997]. Less than full automation is most likely the appropriate level, forcing at least a minimal level of human interaction with automated cognitive tasks. Several researchers have proposed descriptions for varying levels of automation. For example, Sheridan proposed a 10-level scale to describe the decision-making and action execution interactions between humans and automation [Parasuraman, et al., 2000]. This scale ranges in the extremes from decision-making and actions being under full human control to full autonomous control

2.3 Adaptive Automation

For cognitive functions, there does not seem to be one solid solution or method for the design of automation. As Sexton [1988] outlines, the design of automation has many issues to include: to automate or not, how much human-in-the-loop interaction should be required or

permitted, and how intelligent the automated system should be designed. In addition, the appropriate level of interactions between humans and automation can be dynamic or static. The advantages and disadvantages of an automated system for one process may vastly differ from another process, and may vastly differ from one condition to another within the same process. The concept of adaptive or smart automation may provide the best answer.

Adaptive automation responds to the dynamic environment and variations in human performance in choosing the appropriate type and level of automation [Scerbo, 1996]. Adaptive automation is different from the traditional view that automation is either full or not at all. Theorists and researchers began exploring the possible benefits and applications of adaptive automation in the late 1970s with developments in artificial intelligence [Scerbo, 1996]. Adaptive automation had and currently has strong military and commercial aviation applications in assisting pilots with a multitude of cognitive tasks (i.e., decision-making). The ultimate goal for technology with adaptive automation would be to accurately assess the environment, correctly predict and anticipate the user's needs, and most appropriately select the type and amount of cognitive assistance to provide the user, without fail and without information satiation.

While the concept of adaptive automation appears promising, its effective realization for cognitive assistance is a current challenge. A yearlong study by Kaber et al. [2002] attested to the difficulty of designing adaptive automation for cognitive functions. Using Parasuraman et al.'s [2000] four-stage cognitive model for human and automation interactions, Kaber et al. [2002] performed an experiment looking at human performance with the assistance of adaptive automation. They used objective and subjective measurements for performance comparisons between the four cognitive tasks described in the Parasuraman et al. framework. The study concluded that humans found it more difficult to adjust to adaptive automation for cognitive tasks (decision-making) than to adjust to lower-level sensory and psychomotor functions, such as information acquisition and action execution.

2.4 In-Flight Replanner Technology

The stakes are high for both commercial and military pilots in today's flight environments. Battlefields are highly dynamic and uncertain, and are increasingly becoming

more lethal. Commercial traffic is pushing the limits of many airspace structures, and weather hazards do not help. When flight situations do not follow expectations, as they often do not, the cognitive workload for a pilot may increase dramatically. Providing pilots with the best cognitive assistance through decision-support systems is necessary in these time-critical situations. Consequences of poor decision-making are economical at best, and deadly at worst.

In-flight replanner technology integrates essential information elements, using predefined goals and constraints, from the complex and uncertain environment to suggest alternate routes when a conflict or change necessitates a different route trajectory. Such information elements include safety, weather, anti-aircraft threats, terrain, fuel, and traffic for example. While the concept is relatively old, current and future replanner technology push to more effectively and accurately model the demands of the environment and to better assist human cognitive functions. Current computer processing technology is no longer the limitation to modeling and computing alternate routes [Leavitt, 1996]. Instead, the technology is limited by sensor and data acquisition technology to feed the necessary real-time information to the replanner technology, and in large part, by the designer's ability to accurately model the human and automation interactions within this highly dynamic and complex flight environment [Layton, et al., 1994].

In-flight replanners are usually one part of a pilot cognitive assistance system, a system that may include automation for attack planning, survivability planning, reconnaissance planning, and data fusion, for example [Robertson, 2000]. Modern pilot cognitive assistance technology and emerging technology had their roots in the US Air Force Pilot Associates (PA) program of the mid-1980s to early 1990s [Taylor & Reising, 1998]. The goal of the PA was to cognitively assist pilots of military fighter aircraft, giving them the appropriate information when warranted, in the most appropriate manner. While the program initially wanted to demonstrate how artificial intelligence could assist fighter pilots, it actually showed how adaptive automation could be used in complex environments [Scerbo, 1996]. The US Army also had an Army's Advanced Rotorcraft Technology Integration program around the same time, and Lockheed Martin's Mission Reconfigurable Cockpit program followed the PA program [Leavitt, 1996; Dornheim, 1999].

More recently, in-flight replanner technologies have emerged with the US Army's Rotorcraft Pilot's Associate (RPA) program; with Germany's civil aircraft Cognitive Assistance System (CASSY) and Crew Assistant Military programs; and with the French Co-pilote

Electronique military project. The ultimate aim of these cognitive associate or assistant programs can be summarized with Kernstock's [1999] words on the RPA: "[The RPA is] a full-fledged virtual crew member with true cognitive capabilities...designed to substantially increase lethality, survivability and operational tempo".

Modern cognitive assistance programs share two common design philosophies [Taylor & Reising, 1998; Dornheim, 1999; Onken, 1997]. The programs recognized the need for designing an intelligent form of adaptive automation, which was able to vary in response dependent on user needs and environmental demands. In addition, design of cognitive aiding systems should take a human-centered approach, keeping pilots in ultimate command of action execution without information overload. The success of a human-centered design for decision-support automated systems invariably depends on the success of modeling human and automation interactions within naturalistic decision-making environments, such as the dynamic, uncertain, and complex flight environment. There are some key questions for research regarding the human-centered design philosophy. What metrics can be used for determining human performance of cognitive tasks [Kaber, et al., 2002]? In addition, what pilot-vehicle interface best accommodates the pilots in their cognitively demanding environment [Wiener, 1988]?

Let us briefly review some selected literature on recent research efforts to study in-flight replanning from a human factors perspective, using both simulated and actual flight environments. This literature primarily discussed the use of performance criteria to evaluate human performance benefits from using in-flight replanner technology and its associated enabling concepts.

The RPA and CASSY have been flight tested in recent years. From October 1998 to September 1999, the RPA program underwent a series of flight test demonstrations and evaluations, with stunning results [Robertson, 2000; Dornheim, 1999; Colucci, 1999]. For example, pilot survivability using the in-flight replanning automation showed a 75% reduction in combat losses [Colucci, 1999]. Beginning in 1994, Onken [1997] headed flight test studies that evaluated the CASSY's actual flight performance. They tested the CASSY's flight planning and decision-aiding performance, among several other situation assessment and human-computer interface evaluations. They also observed that pilot acceptance of suggested routes was extremely high. The study suggested the need for autonomous planning under extremely time-

constrained situations, as their pilots needed 16 seconds on average to decide on a route proposal.

A pilot-in-the-loop experiment evaluated mission-planning time-critical performance using the Mission Reconfigurable Cockpit automated replanner [Aust, 1996]. The chosen quantitative performance measurement was pilot survivability, and the time-critical situations evaluated were pop-up threats and new target assignments. The study showed several benefits from assisting pilots under time-critical situations with an automated suggested route: pilot survivability increased, replanning response times decreased, and estimated workload reduced. Pilots were told to accept the suggested route, but interestingly, route modifications were still made in more than two-thirds of the trials. The study concluded that automation should provide a generic solution, and allow the pilot to tailor the automated route as desired.

Layton et al. [1994] conducted a study to primarily observe how humans explore data and alternate routes when provided with multiple, computer-generated alternate plans. The study showed that disorientation and missed information were possible results from access to large data sets relevant to the planning process, and warned designers to be careful with the presentation of information. They observed a tendency to rely on the automated route, even a poor route, and suggested that multiple, automated route alternatives may encourage a more global evaluation. The study concluded that humans should be allowed to explore other-than-automated solutions; and that designing a cooperative flight planning system was a significant challenge that needed research into human cognitive processes.

A study conducted by Pritchard [2000] focused on demonstrating benefits from including human decision-making in traditionally fully automated scenarios, such as replanning decisions on reconnaissance missions of unmanned aerial vehicles. The study concluded that with several minutes of time pressure (about five minutes), humans were able to use software to make replanning decisions that accounted for replanning properties that are traditionally difficult to quantify. The researcher suggested that effectively incorporating the human still required research from a human factors and technological perspective. The recommendations recognized the need to study the impact of variable degrees of automation in time-critical in-flight replanning scenarios.

The above studies support the ideas essential to effectively enabling in-flight replanner technology, and cognitive assistance technology overall. Human-in-the-loop experiments are

necessary. Pilots must validate the design of any cockpit cognitive assistance technology, assuring that a true benefit exists before millions of dollars are spent incorporating the technology into actual aircraft. In-flight replanner research needs to focus on time-critical scenarios representative of the real environment. The design of cognitive automation assistance should focus on what is most appropriate and beneficial for the human, which is not necessarily the most quantitatively optimal solution. Current human factors research has the potential to positively benefit the design of cognitive assistance programs for next generation aircraft, such as those for the F-35 Joint Strike Fighter.

2.5 Application to this Research

We performed an experimental study on replanning performance, varying the type and amount of automated information integration and the time pressures. Let us briefly describe how a current naturalistic decision-making model, and a framework for human and automation interactions, fits with our experiment.

Figure 2-1 shows a proposed cognitive model for naturalistic decision-making specifically for in-flight replanning, which was developed by Fan et al. [1998]. The replanning task was broken into four interactive stages: monitor, assess, formulate, and modify. One of the experimental goals was to determine the appropriate automation for processing information and suggesting an initial route when necessary. Using this model, what are the possible interactions with automation?

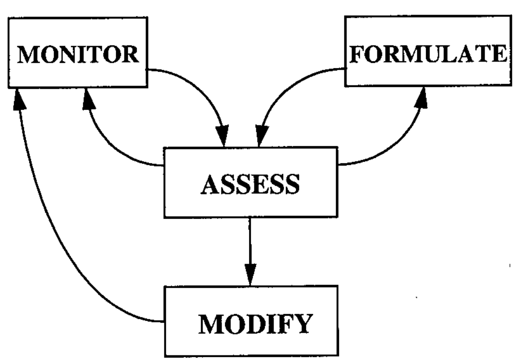


Figure 2-1. Human-Automation Stages for Replanning Task.

Parasuraman et al. [2000] proposed a general four-stage model for systems with automation of information processing, which they developed from their cognitive model of human information processing. Shown in Figure 2-2, the four-stage model included automation opportunities for the information acquisition and analysis input functions, and the decision selection and action execution output functions. Automation level zero indicates no automation, while level 10 indicates a fully automatic function. For the replanning task, we used automation for the information processing input functions (acquisition and analysis) and the decision selection output function. This automation was equivalent to assisting the human for the initial monitor, assess, and formulate stages of the replanning task. The pilot remained in full control at the action execution stage, accepting, modifying, or rejecting the automated suggested solution.

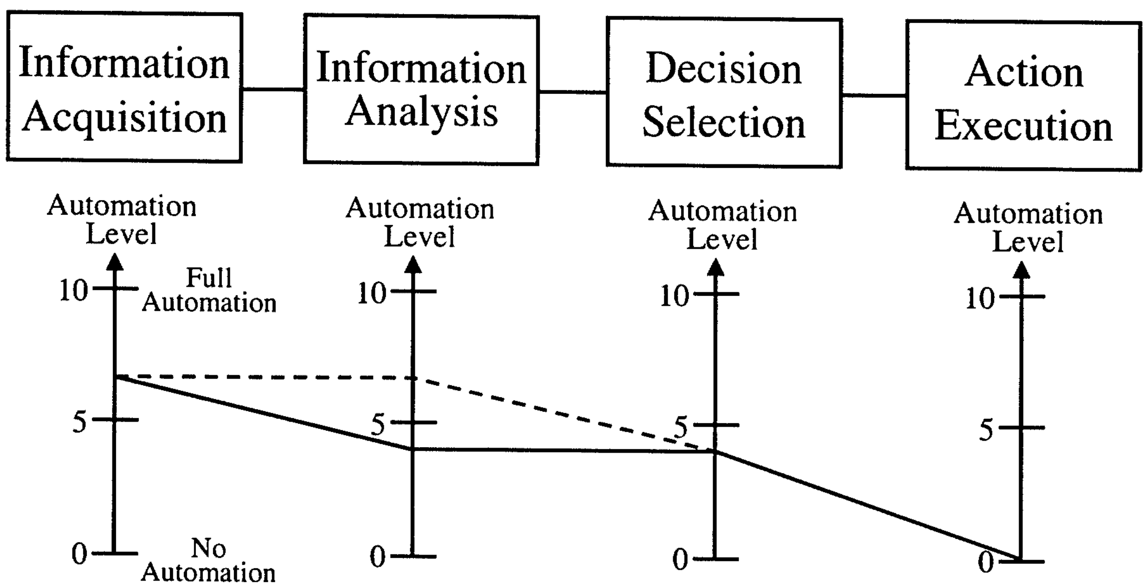


Figure 2-2. Information Processing Four-Stage Model for Automated Systems (adapted from Parasuraman et al. [2000]).

A typical in-flight replanner may only vary its levels of automation at the information analysis stage, otherwise known as adaptive automation. In Figure 2-2, the dotted line represents a higher level of automation than the solid line at the information analysis stage. More automation at this analysis stage may represent the integration and analysis of multiple variables, versus the analysis of a single variable with lower automation. As categorized by Sheridan, automation levels between two and six describe collaborative decision-making and action execution between humans and automation. At the decision selection stage, the in-flight

replanner would then display one suggested route for the pilot, which corresponded to automation level four: "[automation] suggests one alternative" [Parasuraman, et al., 2000].

Please refer to the next chapter, Experimental Design Chapter 2, for more information on the simulated in-flight replanner design. Chapter 2 is the synthesis of concepts and proposed models described in the Background chapter. We further develop the proposed in-flight replanning task model, and describe how the human-in-the-loop experiment objectively and subjectively captured the research interest of decision-aiding automation for information integration, in a time-critical and complex environment.

[Page Intentionally Blank]

3. EXPERIMENTAL DESIGN

Chapter 3 covers in detail each aspect of the experimental design. One of the experimental goals was to quantitatively determine automation and time pressure effects on human decision performance in the in-flight replanning task. We also wanted to objectively define a relationship between information elements of the flight environment, time pressure, automation, and the resulting replanning performance. Standard for any human experiment, we also considered and designed for certain experimental nuisance factors. We determined the nuisance factors to include subjects, map complexity, and learning and fatigue effects. The resulting experimental design was a compromise between available time restrictions and producing meaningful results with a limited resource of easily attainable subjects. What resulted was a repeated measures experimental design using a Graeco-Latin Square test matrix.

3.1 In-Flight Replanning Overview

Overwhelmingly, pilots want to remain in control [Taylor & Reising, 1998]. The design of recent pilot decision-support systems, both actual and experimental, follows this axiom of giving pilots the ultimate decision authority for most decision-making tasks [Aust, 1996; Dornheim, 1999]. Taylor and Reising [1998] suggested the pilots feel increased unpredictability of a process without control of decisions. Specifically for the in-flight replanning task, we modeled the decision-making collaboration between humans and automation after U.S. military in-flight replanner pilot-vehicle interfaces, such as Lockheed Martin's Real-Time Integrated Planner (RTIP) and Rotorcraft Pilot's Associate.

Figure 3-1 illustrates the decision-making model adopted for our in-flight replanning experiment. Both humans and automation observed the information from the flight environment simultaneously. The automation filtered and integrated the specified information, and when triggered by an information update, it then formulated a suggested route. The subject received information inputs from two sources—the flight environment and the automated suggested route. The subject monitored the route and environment in real-time, while simultaneously making

iterative route modifications. When a subject deemed the route satisfactory, whether modifications were made or not, the subject gave the final approval.

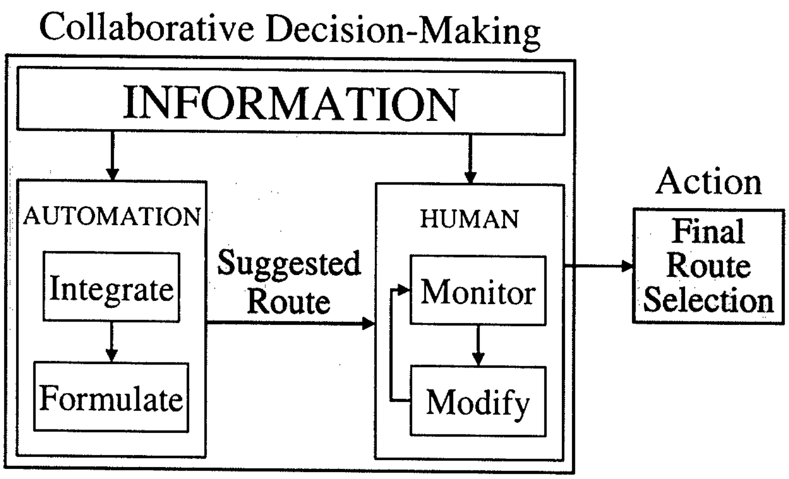
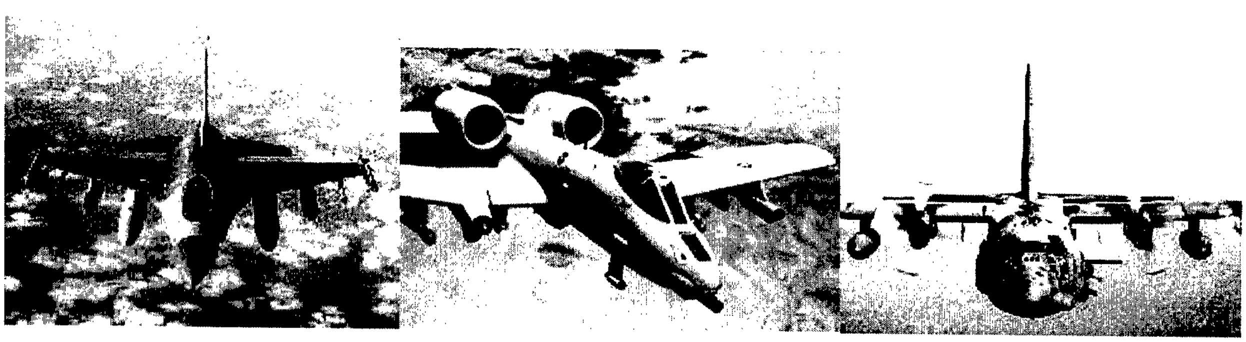


Figure 3-1. In-Flight Replanning Decision-Making Model.

We designed the Dynamic In-Flight Replanner (DIR) experimental simulation using a military combat flight environment theme, and focusing on the in-flight replanning task. This environment was complex, time-critical, highly dynamic, and rich with information input and output sources. The simulated missions were restricted to constant altitude and constant velocity flight of conventional military aircraft with primary functions including air-to-ground, close air support, forward air control, and multi-role. Figure 3-2 shows military aircraft with these functions, which include the F-16 C/D Fighting Falcon, A-10 Thunderbolt II, and the AC-130 H/U Gunship for example [Stringer, 2002].



F-16 **A-10** **AC-130**

Figure 3-2. Aircraft Applicable to Experimental Design (courtesy of www.fas.org).

Figure 3-3 shows the in-flight replanning task from the flight environment to the final route solution. The flight environment included three information elements having first-order effects on the replanning decision-making process: hazard exposure, time-on-target (TOT) requirement, and fuel supply [Leavitt, 1996; Pritchard, 2000; Robertson, 2000]. Furthermore, comments from military pilots insisted cockpit automation must support pilot awareness of status and changes in status of fuel and time-on-target management [Aust, 1996]. Time-on-target is the time at which a pilot arrives at a designated target. A human and automation interaction integrated the information elements to solve a time-constrained problem of developing the best flight plan by the end of the time pressure. The experiment varied the time pressures of the replanning environment, and the types and degree of automated information integration given to the human.

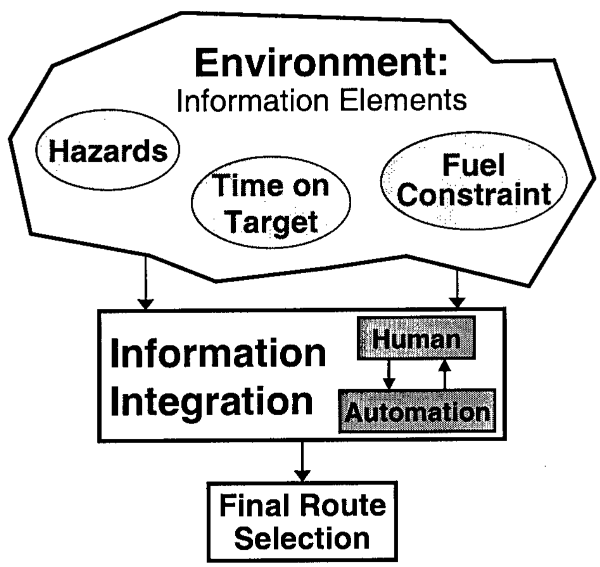


Figure 3-3. In-flight Replanning Task Description.

3.2 Dynamic In-Flight Replanner Design

In general, the experiment involved subjects viewing and interacting with a computer display to develop minimal cost and acceptable routes within a time pressure. Subjects observed a computer monitor flat panel display depicting a Dynamic In-Flight Replanner (DIR). The DIR included a plan-form view of a regional aerial map, which represented approximately 200 square nautical miles (nm). Figure 3-4 shows the entire DIR display with key features labeled: the

navigational map (1), the count-down clock (2), the fuel and TOT constraint gauges (3, 4), the route interaction buttons (5), the appropriate alerts (boxes), and the type of automation route assistance displayed in the upper right.

The aerial map (1) displayed the hazard field, the current and suggested routes, and the associated route points—the start point, a rendezvous point, a target, and an egress (exit) point. The route points, labeled in Figure 3-4, were a green dot representing the start point, a white dot for the rendezvous point, a blue dot with crosshairs for the target, and a red dot for the egress point. The aerial map included threats (labeled), such as terrain, weather, or anti-aircraft missiles, displayed as irregular polygons in four different colors representing four distinct hazard levels. From most to least hazardous, the colors were brown, red, orange, and then yellow. The hazard fields were layered within each other, with the most hazardous level in the center. The brown-centered hazard fields represented terrain, and the red-centered hazards represented threat templates for conventional aircraft. For example, if the pilot flew into a missile threat region, graphically denoted by these ground-referenced threat templates, the pilot was in danger.

Threat templates for missiles include missile track, launch, and intercept templates. A missile track envelope (yellow) includes the largest area where the missile radar can detect and track an aircraft. The launch envelope (orange) includes the area the missile is able to track and launch at the aircraft. The most dangerous is the intercept envelope (red), which is the area a launched missile will most likely intercept the aircraft. In civil and commercial aviation, hazards can be weather, conflicting traffic, or no-fly zones for example, and have similar distinctions in hazard levels.

The TOT requirement also drives the flight plan, requiring the aircraft to arrive at the target at the designated time to ensure mission success; time-on-target does not mean the elapsed time spent flying over a target. In aviation combat, the time-coordinated destruction of a target is critical to mission success in countless scenarios; being early, as well as late, has the potential to cause a mission abort or loss of life. In commercial aviation, the on time arrival at destination is important to customer satisfaction, and ultimately affects company profits. The subject received real-time route TOT information from a horizontal bar gauge (4), indicating the actual flight TOT. The acceptable and unacceptable regions were shaded within the TOT gauge green and red, respectively.

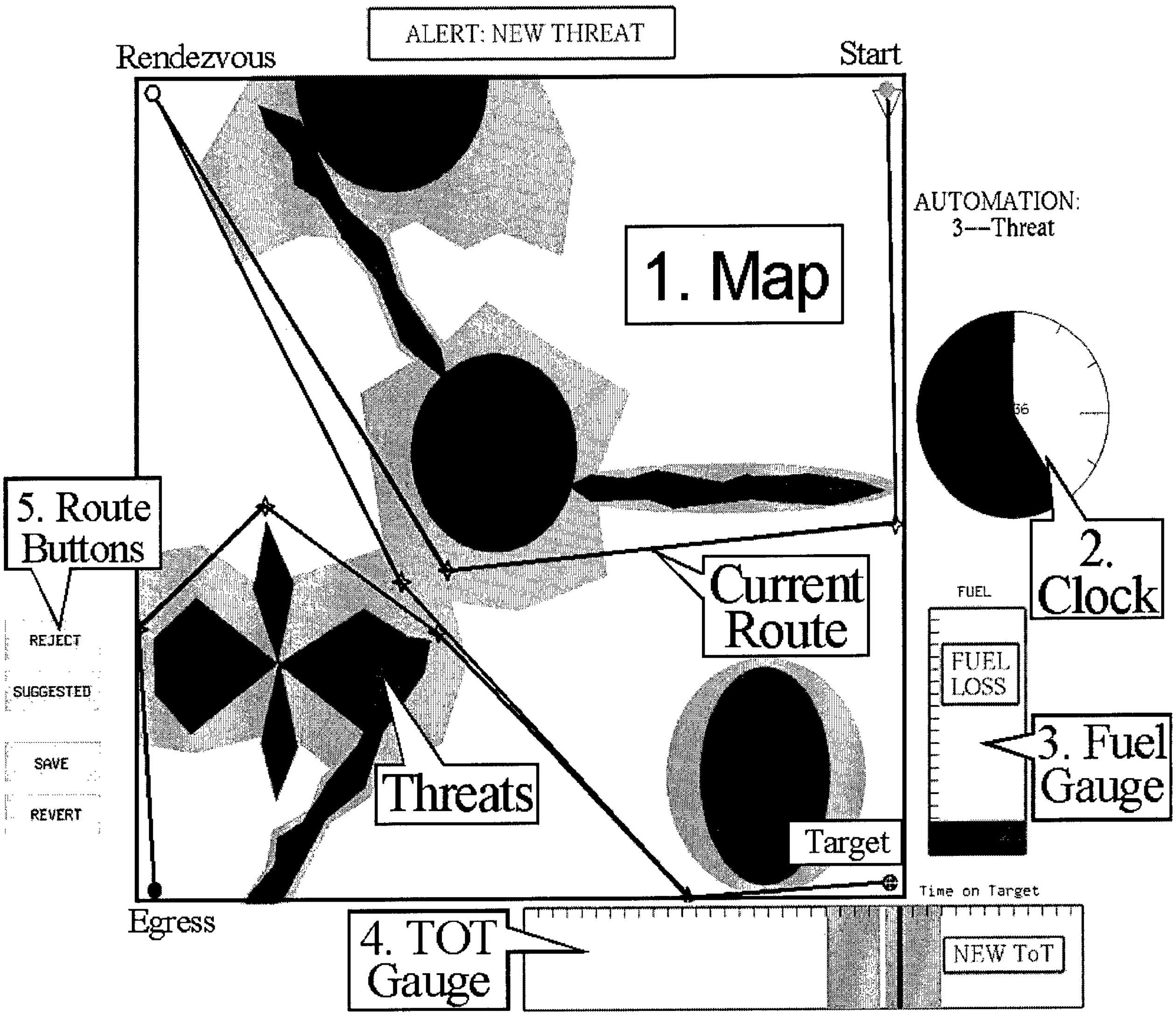


Figure 3-4. Dynamic In-Flight Replanner Interface.

Fuel consumption factors, lastly, heavily influence flight planning. These factors include both economic and life-survival factors, depending on the scenario. In military aviation, the pilot only needs to have enough fuel to complete the mission; while in commercial aviation the economic factors from fuel consumption are of high importance. The subject received real-time route fuel consumption information from a vertical bar gauge (3), indicating the predicted fuel available at the egress point. The fuel gauge turned red when the predicted fuel consumption dropped below the empty fuel tank region. The experiment did not include a cost structure for fuel consumption because military pilots do not factor the cost of fuel into replanning decisions; fuel was simply a constraint.

Using a standard PC mouse, subjects interacted with the DIR to manually develop flight plans under two primary goals: first to ensure mission success, and then to minimize the route cost by reducing hazard exposure and deviation from the TOT goal. Mission success included having fuel to reach the egress point, arriving at the target within an acceptable time window, and avoiding the most dangerous brown-centered hazard level. We were able to artificially impose time pressures using the countdown clock (2), which graphically and digitally displayed the time remaining to replan the current route. The route interaction buttons (5) gave subjects pre-defined gross route modification options (see Section 3.3, Experimental Protocol).

Table 3-1 shows the experimental classifications for each information element condition. We classified the information element conditions based on their influences on the route-replanning task, either mission success or route cost. If a condition of an information element caused a mission failure, that condition was a constraint. We labeled conditions causing mission failure probabilities less than one as soft costs, incurring route cost penalties for these violations.

Table 3-1. Information Element Classification.

Information Elements		**Classification**	
		Route Cost	**Mission Constraint**
Hazard (local)	Brown		X
	Red, Orange, Yellow	X	
Time-on-Target (global)	Deviations	X	
	Max Deviation		X
Fuel (global)			X

The red, orange, and yellow hazard fields represented threats with some probability of a mission failure less then one. The probability of mission failure from a threat was linearly proportional to length of exposed route and to the threat level. A subject's TOT could be either a soft cost or a constraint. As long as the subject had a TOT within a determined acceptable time window, TOT remained a soft cost. The mission failed if the subject arrived at the target before or after the time window around the TOT goal; this TOT condition was a constraint. Lastly, fuel was only a constraint, forcing subjects to have enough fuel to safely reach the egress point.

We also classified the information elements into local and global information, which was based on the manner a human processes the information. For example, a subject could reduce route cost simply by making appropriate local route adjustments to avoid local hazards. A subject must always monitor global effects, however, even when processing local information. Both TOT and fuel were global information elements. For example, any route modifications changed the projected fuel levels, even when modifying the route to avoid hazards.

Lastly, including the TOT requirement and fuel constraint information into the replanning decision process was an important experimental design factor. In-flight replanning technology has recently been able to approximate the least cost and constrained route solution in a predictable amount of time; this is an NP-complete problem otherwise known to be computationally intractable [Garey & Johnson, 1979]. Therefore, now is the time to take a human factors perspective into the design process of in-flight replanners: How can designers best present the constraint information to pilots in the cognitively demanding flight environment? We wanted to better understand how the design of automation that integrated TOT and fuel information would affect human behavior in the replanning task, and to develop its relationship with other information elements.

3.3 Experimental Protocol

Subjects first viewed an initial pre-flight plan containing the original route, Figure 3-5 (Previewed Mission). The original route was a white dashed line, which was connected in serial order from the start point to the rendezvous point, the target, and finally the egress point. The pre-flight plan had a good original route, ensuring mission success with the latest available pre-flight information on hazards, and fuel and TOT restrictions. Subjects could take as long as desired in viewing the pre-flight plan, giving them a sense of the mission. Subjects used the pre-flight plan to become familiar with currently known threat locations and severity levels, and where the start, rendezvous, target, and egress points were.

After becoming familiar with the pre-flight plan, the data collection and time-pressured portion began when the subject indicated being ready. The start of data collection also triggered sudden changes in the environment and updates to the mission restrictions, Figure 3-5 (Updated Mission). An environmental change included new popup threats. Mission restriction updates

included a new TOT goal, a new TOT acceptable window width, or a new fuel restriction due to dumping fuel, for example. Environmental or mission restriction changes did not occur again for the remainder of the time pressure, which the countdown clock imposed. We were interested in subject performance in response to a one-time change, in solving a static problem; having multiple updates within one mission would not add valuable information for our research focus.

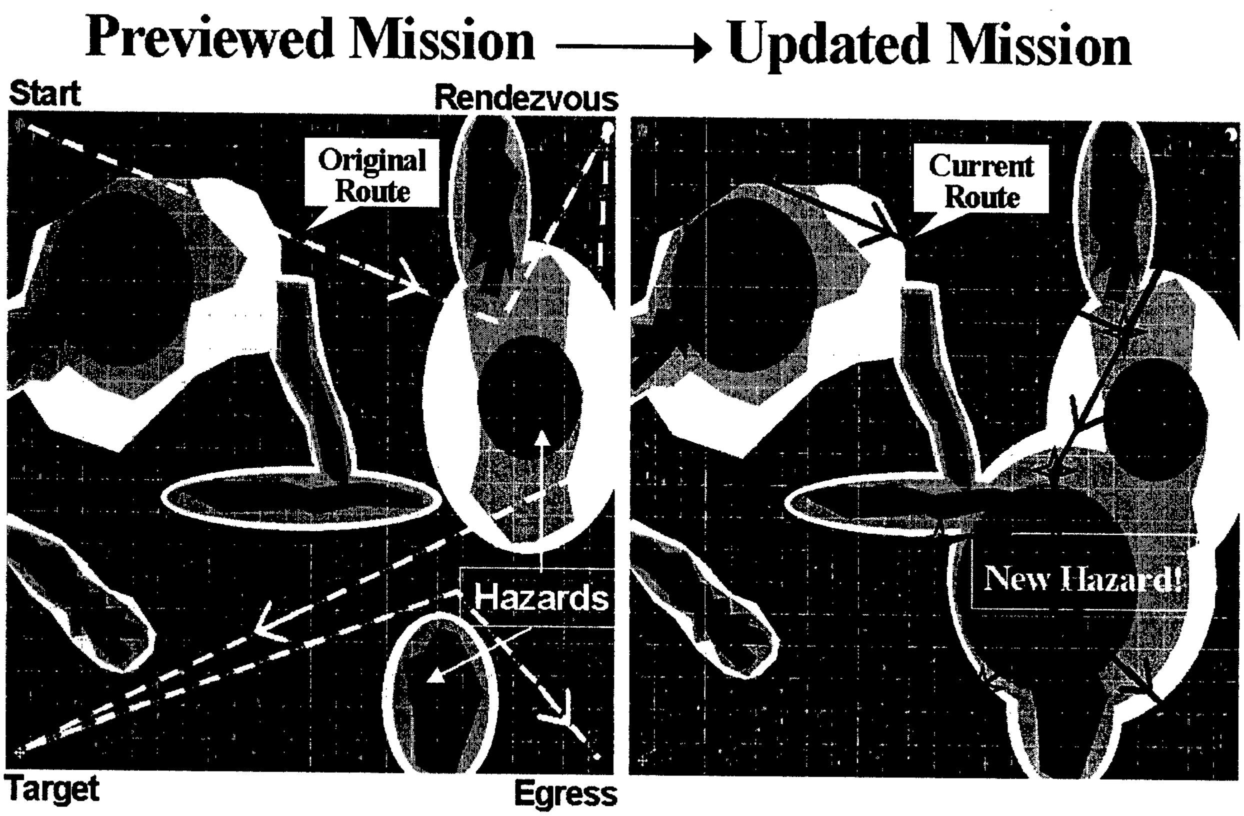


Figure 3-5. Dynamic In-Flight Replanner Map Update.

(arrows annotate route direction, not in actual display)

Subjects received one of three types of automation assistance in the form of a suggested route when a change in environment or mission restrictions occurred (see Section 3.5, Independent Variables). The automation only suggested an initial route, and did not update or respond in real-time to the subject's route modifications. The automation processed the environment's information elements according to its type, and displayed a suggested route as a magenta-colored and dashed line to the subject. A solid blue line connected by blue star-shaped waypoints represented the current and modifiable route, which initially mirrored the automated

route suggestion if there was one. Under time pressure, subjects used a PC mouse to manually replan the current route trying to achieve an acceptable and minimal cost route. For reference, the suggested route always remained displayed in the background. To avoid unnecessary display clutter, the DIR did not display the original route when automation suggested a route.

Subjects modified the current route by moving, and adding or deleting route waypoints. Moving waypoints, or "rubber-banding," used the mouse-click-and-drag method. By mouse clicking on a waypoint or specific route location, the subject could add (left button) or delete (right button) waypoints as desired. In addition, the DIR had several pre-defined route modification options. Subjects could reject the current route, which cleared all the waypoints and left the route connected by only the start, rendezvous, target, and egress points. The subject could also save the current route for later reverting to the saved route state. This function allowed subjects to safely explore various solution spaces with the option to revert to a good saved route when getting low on available time or when making the route worse.

Figure 3-6 summarizes the experimental protocol temporally, where increasing time is to the right. The subject first previewed indefinitely the mission to get an overview of the hazards, the route, and the constraints. When ready, the subject started the scenario (A). At this point, there were updates to the mission requirements or changes within the actual environment, and the automation suggested a route based on its assistance type. The subject manually replanned the current route until the expiration of the time pressure.

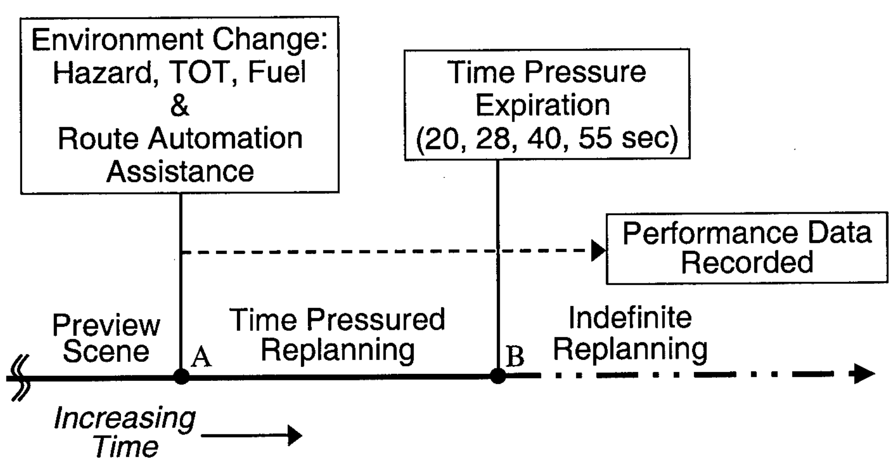


Figure 3-6. Experimental Scenario Timeline.

At the time pressure expiration (B), a feedback screen appeared displaying the status of each mission constraint. For example, "Fuel: Not Satisfied" would display in red letters if a subject failed to meet the fuel constraint. Four of the sixteen scenarios continued after the feedback screen disappeared, allowing subjects to replan indefinitely until satisfied with their minimal cost route. As described in Section 3.5 (Independent Variables), these scenarios determined a subject's optimal replanning performance. The DIR software also recorded performance data, to include route cost and constraint violations, after each route waypoint modification and at the expiration of the time pressure.

Before the experiment, subjects filled out a standard human experiment consent form and a short biographical data sheet. During the experiment, there was a brief questionnaire after each data collection scenario. There was also a post-experiment short interview with the experimental proctor discussing issues related to training, replanning performance, and automation assistance. The experimental consent form and questionnaire can be found in Appendix A.

3.4 Human Factors in Display Design

The DIR display used many human factors display design principles and guidelines. In addition, the DIR display format and subject interactions were modeled in part from Lockheed Martin's Real-Time Integrated Planner (RTIP). The primary display design principle used was the compatibility of proximity principle. This principle states decision-making tasks will benefit from the close display proximity of similar information, whereas close proximity of information will hinder different tasks that require independent processing or focused attention [Sanders & McCormick, 1993 (Chapter 5)]. The placement of constraint gauges and the countdown clock were in close display proximity to each other because this information was critical to the replanning decision-making process. The route interaction buttons in middle-left of the display were in close proximity to each other, and far from the constraint information used for information processing.

The constraint gauges used moving pointers against a fixed scale, which humans generally prefer [Sanders & McCormick, 1993 (Chapter 5)]. The scales followed common perceptions of pointer movements for indicating increasing and decreasing quantities. The pointer on the vertical fuel gauge moved lower as the fuel decreased, eventually into the empty

region at the bottom of the gauge. The pointer on the horizontal TOT gauge moved right as the flight time to target increased. The constraint gauges were also qualitative scales coded by colors to indicate to the subject approximate conditions [Sanders & McCormick, 1993 (Chapter 5)]. When in normal conditions for fuel and TOT, the gauge was green. When in unacceptable conditions, the gauges turned red indicating danger or warning.

The map, at the display's center, closely paralleled Lockheed Martin's RTIP. However, we had to be careful with too much clutter in providing the subject with visual information—the threat templates and route structures. The RTIP design showed the original, suggested, and current routes on the same display. We only displayed the suggested and current routes to reduce clutter without taking away important information. As with the RTIP, the current route was blue, and the suggested route was magenta. The threat template colors were (in decreasing severity) red, orange, and yellow, which represented the same threat severity descriptions used for the RTIP.

For the DIR simulation interface design, we incorporated human factors considerations, as well as considering current in-flight replanner displays. This fostered information processing and integration for the decision-making task, and expedited the learning of DIR interactions when training. Ultimately, we could more effectively and accurately measure the independent variable effects on the dependent variable with a display interface designed with human factors considerations.

3.5 Independent Variables

This study included two controlled independent variables of interest: automation assistance category and time pressure.

Automation Assistance

Automation assistance came in the form of a suggested route displayed to the subject when a scenario update occurred due to new hazard information or changes in TOT or fuel requirements. The automation provided only a one-time route suggestion, and did not update in real-time in response to subject inputs or route constraint violations. Control of final route

acceptance always remained with the subject. Labeled by the information processed, three automation assistance categories were evaluated: None (the control group), Partial, and Full automation assistance. Two subcategories further divided Partial automation assistance into Hazard and Constraint automation assistance. For brevity, the rest of this document refers to the automation assistance categories as None, Partial (either Hazard or Constraint), and Full. In all scenarios with None or Partial, the automated route suggestions were unacceptable, violating at least one mission constraint. Table 3-2 describes the relationship between automation assistance and the corresponding information elements it processed.

Table 3-2. Automation Assistance Description by Information Element.

Automation Assistance	**Information Element**		
	Fuel restriction	TOT restriction	Hazards
None			
Partial – Constraint	X	X	
Partial – Hazard			X
Full	X	X	X

None, Partial, and Full represented a hierarchy in levels of automated integration. We chose the Partial categories to analyze various human factors implications regarding human decision-making performance, such as the effects of different human cognitive workload modalities and human information processing theories. Accounting for hazards in route replanning, for example, was primarily a visual workload task requiring the processing of local information. Whereas, considering fuel and TOT mission constraints required integrative processing of global information, and was primarily a mental workload task. Detailed discussions of each automation assistance category follow:

1. None

There was no route automation assistance with None. We used None as the control group, the reference for subject replanning performance with route planning automation assistance. Without any automation assistance, the subject would replan unaided by modifying the original route as necessary in response to a mission update. All scenarios with None had

initially unacceptable routes, violating at least one mission constraint after the information update.

2. Partial

a. Constraint

Constraint processed the scenario's global constraint information only, satisfying the scenario TOT and fuel requirements. Constraint suggested a route that minimized the TOT cost, while still meeting the fuel constraint for the given scenario. Constraint automation achieved this by relaxing the original route until the TOT cost was minimized and there was enough fuel, without regard to environmental hazards. In the case that the original route arrived at the target too early, the automation inserted a holding pattern into the route at or near the route starting point. Subjects needed to initially process and integrate the scenario's hazard information with the Constraint suggested route. In all scenarios with Constraint, the suggested route satisfied the TOT and fuel mission constraints, but did not meet the brown-hazard constraint. Therefore, subjects manually replanned to locally avoid hazards and to lower route costs.

b. Hazard

Hazard processed the scenario's hazard information only. Hazard suggested a route that locally minimized the route's exposure to hazards, avoiding hazards completely when possible. Hazard locally relaxed the original route where necessary to avoid hazards, without regard to the other mission constraints. Hazard was only able to detect the highest two of four hazard levels (brown and red), leaving the lowest hazard levels for the subject to consider. This represented automation not receiving or being able to process all the information. Subjects needed to integrate the suggested route with the scenario's TOT and fuel constraints and the lower-level hazard information. Starting with a low cost suggested route, subjects primarily needed to assure the route satisfied global mission constraints by the

end of the time pressure. In all scenarios with Hazard, the route suggestion did not meet either the fuel or TOT constraint.

3. Full

Full integrated both hazard and constraint information simultaneously, suggesting a route that minimized local hazard exposures and satisfied the global fuel and TOT mission constraints. Full first activated the Hazard component, minimizing route cost. Then, Full activated the Constraint component to force an acceptable route by relaxing the Hazard minimum cost suggested route. While this automation produced an acceptable and low cost route, there was still a significant amount of cost improvement possible. Subjects generally modified the Full suggested route if they felt improvement on the suggested route was possible under the time pressure.

Time Pressure

From an earlier study of in-flight replanning decision aids, Fan, et al. [1998] suggested time-critical events focused on safety without regard to efficiency and were on the order of a few seconds. They also defined tactical replanning as having the time for safety and efficiency concerns, and as taking a few minutes. Our study tried to find the transition point from time-critical to tactical replanning, when subject motivations shifted from primarily safety to both safety and efficiency.

We chose four time pressures based on pilot studies and time-critical interests: 20, 28, 40 and 55 seconds. Having hypothesized that the greatest performance changes occurred at the highest time pressures, we chose the times from a logarithmic scale between 20 and 55 seconds to best capture this performance transition.

While all experimental scenarios imposed an initial time pressure, four scenarios allowed a subject to replan indefinitely after the allotted time expired. Without time pressures on replanning, we captured a subject's optimal replanning performance through the generation of their best routes. The analyses used the data without time pressure as a reference for the data with time pressure and for between-subject performance comparisons.

3.6 Dependent Variables

Subject replanning performance was the primary dependent variable. We used several objective and subjective measures of replanning performance. The primary objective measurement was the route cost at the end of a time pressure. Route cost provided a detailed and convincing measurement for a subject's replanning performance. There were several supporting and less complete objective measurements for subject performance: the type and number of mission failures and the number of route modifications. A questionnaire and post-experimental interview were the two subjective measures of replanning performance.

Route cost was the primary quantitative measure of replanning performance, which was a transformed value of the raw route cost from the experimental data output. The raw route cost was a function of route hazard exposure and deviations from the TOT assignment. Following is the equation used to calculate the raw route cost.

$$\text{Raw Route Cost} = A \sum_{i=1}^{\#\text{Colors}} \left(Length_{\text{Route}_i} \times Cost_{\text{Color}_i} \right) + B \left(\exp^{\left(b_1 \times \left| \frac{t}{t_0} \right| \right)} - 1 \right)$$

Subject performance directly influenced the route length through any hazard and the time deviation from the TOT goal. Hazard exposure cost was linear with the length of the route intersecting a hazard. The hazard cost ratio of {red : orange : yellow} was equal to {10 : 3 : 1}. While brown hazards were constraints, there was an additional route cost penalty for intercepting a brown hazard. Time-on-target cost had an exponential increase up to the maximum allowable TOT deviation, which was defined by the acceptable time window before and after the TOT goal. The t/t_o ratio represented the normalized TOT goal deviation, where t was the actual TOT deviation and t_o was the acceptable TOT deviation. There were no additional route cost penalties for exceeding the maximum acceptable TOT goal deviation, it was simply a mission constraint. In addition, a cost model was not fit for conserving fuel. As discussed previously, fuel was only a constraint in military flight scenarios.

Based on several pilot studies, we adjusted the variables A, B, $Cost_{Color}$, and b_1 to appropriately motivate the subjects in following the goals of a representative in-flight replanning scenario. In all scenarios, the variables were adjusted to force a balance between competing in-

flight replanning goals of minimizing route hazard exposure, and meeting the TOT and fuel requirements. Table 3-3 lists the actual values used.

Table 3-3. Raw Route Cost Structure.

Variable	Value
A	50
B	10,000
Red Cost	1
Orange Cost	0.3
Yellow Cost	0.1
b_1	2

Figure 3-7 shows the relative raw route cost structure for TOT and hazards costs in relation to flight time. This plot gives a good graphical indication of the cost motivations from TOT and hazards felt by the subjects. To enhance the physical meaning of the raw route cost, we referenced the cost relationship to a tangible unit, flight time in minutes. The vertical axis represents the raw route costs, and the horizontal axis represents flight time in minutes. We modeled this time relationship after a military flight scenario example, using aerial map dimensions of 200 by 200 nm, and an assumed typical combat velocity of 800 nm per hour.

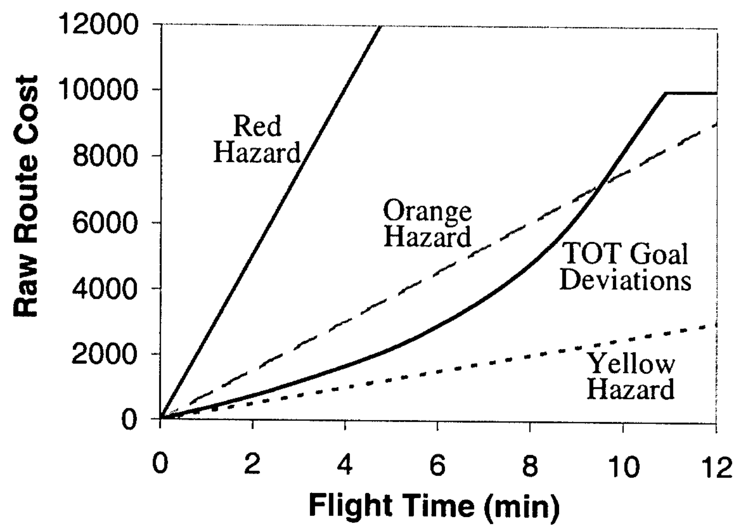


Figure 3-7. Raw Route Cost Versus Flight Time Relationships.

For example, a red hazard raw route cost of 10,130 was equivalent to flying four minutes through only red hazards. Alternatively, an orange hazard cost of 3040 or a yellow hazard cost

of 1010 was also equivalent to flying four minutes through only the respective hazard color. Four minutes of deviation from the TOT goal was equivalent to a cost of 1637. The maximum TOT deviation raw route cost was 10,000, which occurred at a TOT deviation equivalent to plus or minus 10.9 minutes of flight time.

In addition, the DIR included an invisible cost penalty buffer surrounding each hazard level, which acted as the hazard safety buffer. The buffer had a cost penalty proportional to the hazard level, and had a constant width independent of the hazard level. Humans naturally place a safety buffer on most decisions, with some humans being more conservative than others. However, the artificial buffer forced all subjects to maintain a minimum separation from hazards, or receive a cost penalty. We hoped the buffer would better motivate subjects to stay clear of the more severe hazards, particularly the brown and red levels.

3.7 Human Effect

Inherently, each subject varied in their ability to make replanning decisions. The experimental design and data analyses had to take into account these inherent differences between subjects; otherwise, the observed measurement values would not accurately reflect the independent variable effects, such as time pressure and automation assistance effects on route cost. The experiment and data reduction used several methods to reduce variability between subjects.

The subjects first needed to have similar decision-making models and performance incentives for making valid between-subjects comparisons on replanning performance. A detailed and comprehensive training tutorial ensured similar decision-making between subjects. Furthermore, the tutorial trained each subject to understand completely the mission environment and goals, and to interact with the DIR at the same skill level.

Ensuring similar performance incentives between subjects was another experimental issue. It would be difficult to compare the performance results between a motivated and lazy subject. Many commented on the video game –like nature of the experiment, which augmented the subject's already inherent desire to perform best. To the best of our knowledge, the subjects chosen for the experiment and included in the data analyses were inherently motivated to perform well under pressure.

Despite the aforementioned efforts to reduce performance variations between subjects, individual subject effects innately occurred. While the statistical analyses model the human as an effect, we manipulated the data to help reduce between subject variations. We decided to normalize each subject's data by his or her own optimal performance, rather than use an absolute reference. Thus, optimal performance for each subject had the same route cost value after normalization.

3.8 Map Effect and Map Complexity Design

The experiment used four baseline maps to generate 16 effective data collection scenarios needed for the Graeco-Latin Square four-by-four test matrix. Unavoidably, subject performance was highly dependent on the map difficulty and complexity. Therefore, we created each of the base maps following the same iterative process to help assure the maps were of similar complexity. This allowed us to more accurately analyze the data with respect to automation assistance and time pressure effects.

Using an iterative process, we created four baseline maps of similar cognitive and quantitative complexity. We designed each map to force a balance in goals between fuel constraint, TOT deviation, and minimal hazard exposure. The minimal cost routes, therefore, did not allow any of the individual information elements to be at their optimal states. For example, the route could never completely avoid all hazards without running out of fuel, or the route could not reach the target exactly on time without being exposed to hazards.

There were two main factors for the design of map complexity: cognitive and quantitative complexity. Cognitive complexity referred to the subject's perceived workload for each map in solving for the minimal cost route. For example, threat field density, threat placements, and retrictiveness of constraints influenced cognitive complexity. The quantitative complexity was the actual route cost due to hazard exposures and deviations from the TOT goal. The minimal cost route for each map, or optimal solutions, needed to be similar in cost; otherwise, a quantitative analysis on route cost would not be valid.

In general, we first created the maps graphically and then iterated slightly until the quantitative values for constraint restrictions and minimal cost routes were similar to within a tolerable limit. First, we forced subjects to consider the same solution space by using the entire

map region for each scenario. This was accomplished by placing each route point—start, rendezvous, target, and egress points—in the map corners in the same order. To avoid further complexity issues, the map solution space did not go beyond the displayed boundaries.

We then designed similar hazard fields for each map. The strategic placement of brown-centered hazards forced a decision between at most two possible solutions at each route point. We placed two or three red-centered threats along the intended route before the target. The same amounts of threats were placed between the target and egress points. The spatial size of each hazard was another design issue. For example, having four spatially large threats would be different from having four spatially small threats. Each map's hazard field covered nearly the same spatial area.

The iterative design process between route design and mission constraints was next. The optimal route for each map, incorporating the fuel and TOT constraints, needed to result in similar raw route costs. First, we found an initial low cost route due only to hazard exposures for each of the four maps. Next, we artificially adjusted the TOT and fuel restrictions around this low cost route to be the same for each map. We tried to design the constraints to be restrictive enough on the initial iteration to render the initially guessed route the lowest cost route possible for that specific map. The Scenario Editor software showed in real-time the effects from route and constraint adjustments, which significantly reduced the map generation time.

After the initial iteration in design of route and constraint values, we then checked the route to assure no other acceptable routes had a lower cost within the entire map solution space. When a lower cost route was found, we iterated the fuel and TOT constraint adjustment process with a different low cost route. The minimal cost routes for each map needed to have similar costs to within acceptable limits. Figure 3-8 shows the final design of the four maps used in the experiment, with the associated optimal routes. The minimal raw route costs for each map were 15940, 15490, 15010, and 16000, which we considered very similar. Although not shown in Figure 3-8, the fuel and TOT restrictions were the same for each map.

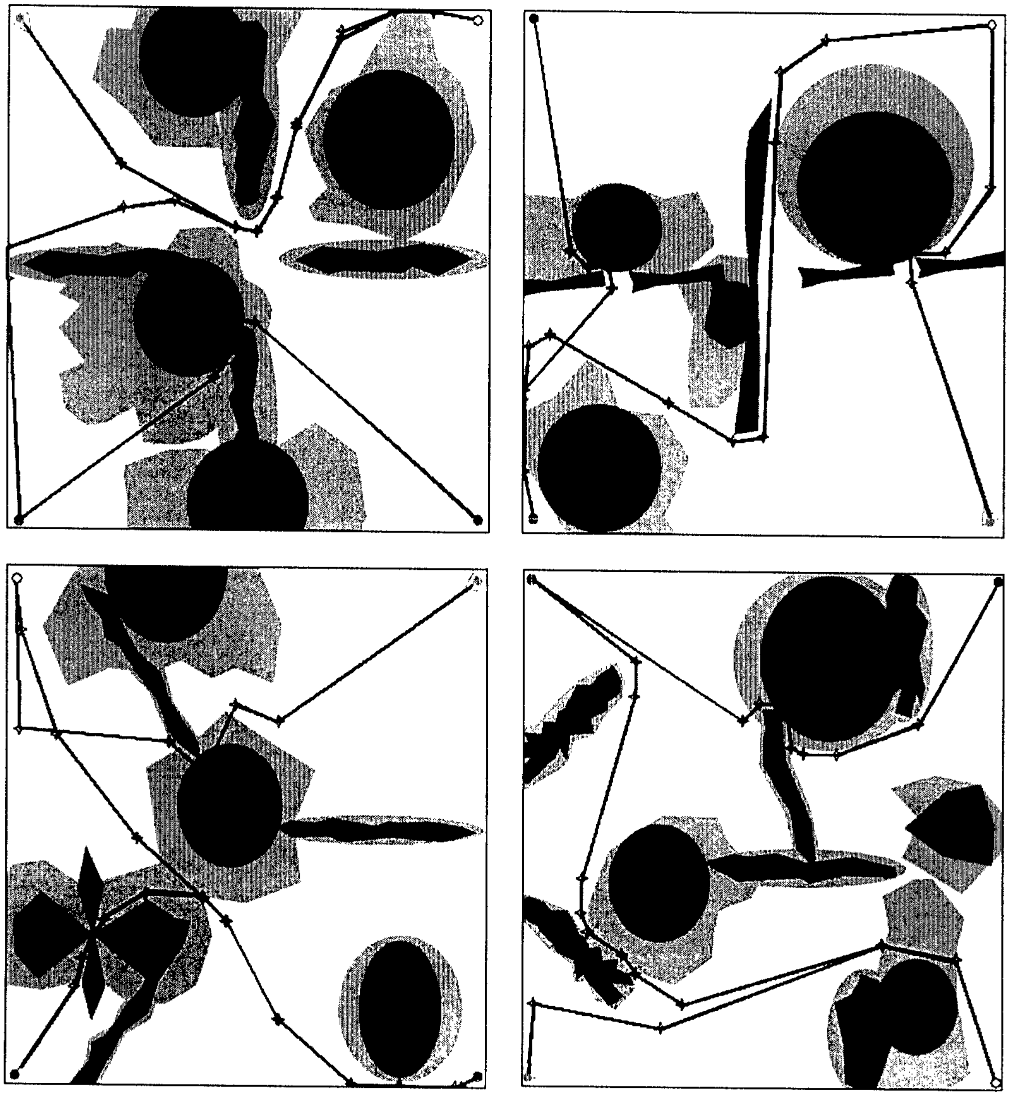


Figure 3-8. Maps with Optimal Routes.

While we used only four baseline maps for 16 scenarios, subjects were not able to perceptually distinguish them. Map rotations and different preview maps helped ensure subjects would not identify similar maps. In the experiment, the four baseline maps were the updated maps. The baseline maps then had hazards removed and route restrictions relaxed to produce the original previewed maps. Furthermore, we rotated each map four times by 90 degrees as illustrated in Figure 3-9. In essence, there were 16 effective maps of similar cognitive and quantitative complexity, yet each was perceptually different to the subjects.

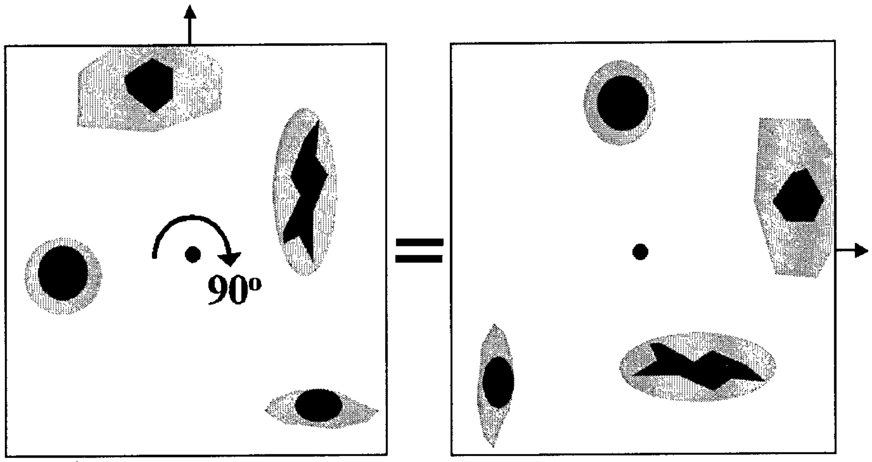


Figure 3-9. Map Rotation Example.

3.9 Test Conditions

We performed a multivariate repeated measures experiment, with each subject experiencing all test conditions in the same trial order. A repeated measure design is a within subjects block design, in essence using each subject as their own control group. The advantage of a repeated measures design was to have a more accurate data analysis because one does not have to average main effects between subjects. However, this was at the expense of longer experiments per subject since each subject needed to run all test conditions. We concluded, however, that the benefit far outweighed the extra time per subject penalty.

Shown in Figure 3-10, the repeated measures experiment followed a four by four Graeco-Latin Square test matrix design with four main factors: time pressure, automation assistance, map difficulty, and map rotation. The Graeco-Latin Square design exposes a subject four times to each group within each of the four main factors, all within 16 trials. The experiment used four base scenario maps of similar complexity, each rotated four times by 90 degrees to effectively achieve the 16 perceptually different scenarios. Because the map reference frame was arbitrary, there was not a direct correlation with a 90 degrees rotation of one map to another, and findings from rotation analyses were inconclusive.

Figure 3-10 also shows the trial order for each time pressure and automation assistance condition. The Graeco-Latin Square allowed consecutive scenarios to have a different map, time pressure, and automation assistance than the previous scenario. Trials 13 to 16 have asterisks denoting the optimal replanning scenarios, where replanning continued indefinitely after the time

pressure expired. The results from the optimal performance runs gave us more data points to better analyze and reference performance under time pressures. These four scenarios captured a subject's optimal replanning performance by allowing subjects to replan indefinitely after the original time pressure expiration.

Automation Assistance \ Time Pressure (sec)	20	28	40	55
None	Map 3-rot 2 Trial 5	Map 4-rot 1 Trial 2	Map 2-rot 4 Trial 16*	Map 1-rot 3 Trial 11
Constraint	Map 4-rot 4 Trial 14*	Map 3-rot 3 Trial 9	Map1-rot 2 Trial 7	Map 2-rot 1 Trial 4
Hazard	Map 2-rot 3 Trial 12	Map 1-rot 4 Trial 15*	Map 3-rot 1 Trial 1	Map 4-rot 2 Trial 6
Full	Map 1-rot 1 Trial 3	Map 2-rot 2 Trial 8	Map 4-rot 3 Trial 10	Map 3-rot 4 Trial 13*

Figure 3-10. Graeco-Latin Square Test-Matrix.

3.10 Training

Subjects extensively trained on the DIR training module before starting the data collection scenarios, training for an average of 3 hours and 10 minutes. A comprehensive tutorial, along with a verbal instruction portion, guided each subject through 12 training scenarios. The training focused on becoming familiar with the DIR display interface and proficient at the replanning task. Subjects learned and implemented strategies to best use route automation assistance under the time pressures. Most importantly, the training ingrained into the subject's decision-making model the flight mission goals and restrictions; as well as the route cost heuristics for exposures to different hazard levels and deviations from the TOT goal.

The training scenarios first introduced each automation route assistance category at the lowest time pressures, giving subjects at least 55 seconds to replan. These initial training scenarios also allowed subjects to continue replanning after the time allotment expired. The last six training scenarios introduced subjects to the extreme time pressures and environment difficulties representative of the actual data collection runs. The final training scenarios ended at

20 and 28 seconds, without the opportunity to replan indefinitely. Without being able to continue replanning upon the expiration of time pressure, subjects better experienced the feelings of time pressure in having to quickly complete an acceptable and low cost route.

There were several features in the DIR training module not used in the actual data collection runs. The training module displayed separately and in real-time the route cost components due to hazard exposures and TOT goal deviations. The route cost displays were two identical vertical bar gauges, with the cost digitally displayed below the gauges. With real-time feedback, subjects were able to quickly learn the cost heuristics and develop strategies for route replanning. For example, one replanning strategy subjects learned using the route cost displays was how close the route could be to a hazard before intersecting its constant-width safety buffer, which the DIR did not display. In addition, the training module allowed subjects to repeat any practice scenario an unlimited amount of times.

For a more detailed description of the experimental protocol, subject interactions, and display interfaces, reference the actual experimental training tutorial in Appendix B.

3.11 Software Development

The experimental study used software developed in Microsoft's Visual C++ environment, using C and OpenGL programming languages. We chose these languages for several reasons. C and OpenGL are widely used and accepted in academics and industry, and consequently have a large foundation of literature and on-line support. C and OpenGL are robust enough to allow for easy programming of real-time simulations. Both languages are highly portable. Lastly, we had prior knowledge and experience with programming in C and OpenGL.

The OpenGL library, along with the OpenGL Utility Toolkit (GLUT), is highly portable because it does not contain any window system-dependent operations. We used GLUT as the programming interface for OpenGL, since OpenGL does not contain any window system operations [Kilgard, 1996]. GLUT was simple to use and understand, and it still had all the functionality needed for our experimental purposes.

Figure 3-11 illustrates the relationship between the two primary software programs used in the experimental study. The designer used the Scenario Editor (SE) to develop the simulated

missions, which the Dynamic In-Flight Replanner (DIR) needed to function. The DIR outputted the raw performance data used in the data analyses.

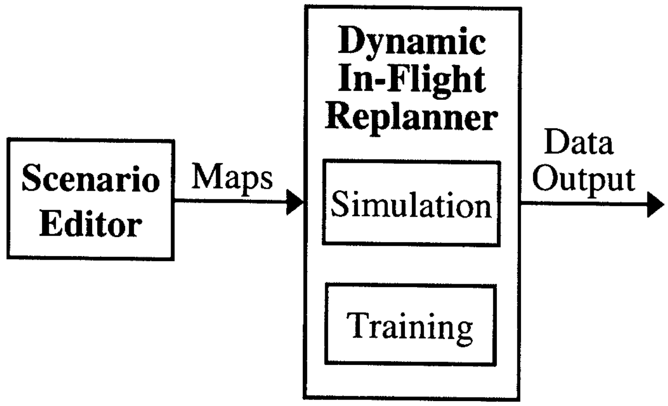


Figure 3-11. Software Interactions Flow Chart.

The Dynamic In-Flight Replanner (DIR) software consisted of two main modules: the part-task in-flight replanning simulation and the training module. The training module's purpose was to methodically teach each subject the display format and proper interactions with the in-flight replanner. Subjects ran the actual experiment on the DIR part-task simulation. The replanner was dynamic in the sense that relevant information updated real-time in response to subject inputs. Functionality was nearly identical in both modules, with a few more functions in the training tutorial. Because the DIR software required real-time user interaction capability, the experiment only used computers able to process in real-time the demands for graphics rendering. The mouse was the only DIR input device.

The DIR software also included a data-logging module, used in parallel with the simulation and training. This module allowed for the easy manipulation of the data output format, to include choosing what salient data to record, and saved valuable time in data reduction. It recorded data after each subject DIR interaction, and at the beginning and end of each scenario. The DIR software provided for a low-cost, easily accessible and repeatable, and modifiable human experiment.

The Scenario Editor software complimented the DIR simulation software. The Editor provided the environment to completely develop and edit the simulation maps and the route constraint parameters as needed. It was primarily a graphics editor paint program. The user could draw polygons, lines, and points in most of the primary colors, as well as design the route.

The editing functions included delete, copy, move, clear screen, undo, flip horizontally or vertically, and edit polygon points to name a few. The Editor also had the capability of displaying real-time the effects of manipulating route fuel and TOT constraint variables as desired. Having control of these variables allowed for the generation of map and route restrictions of similar complexity—previously discussed in Section 3.8, Map Effect and Map Complexity Design. The editor used both the keyboard and mouse as input devices. The editor also had a file input output module for saving or loading map files. The DIR simulation required these output files to run.

For a complete and detailed technical description of both the DIR and Editor programs, please refer to Appendix C: Software Architecture.

[Page Intentionally Blank]

4. QUANTITATIVE RESULTS

Chapter 4 describes in detail the main effects on route cost due to time pressure and automation assistance, and the associated interaction effects. Then the findings supporting the concept of a limited temporal benefit from automation assistance are discussed. Finally, the chapter covers the supporting and additional quantitative findings from analyses of mission failures, route modification events, and Full automation assistance.

4.1 Subject Demographics

The experiment was conducted in-house at the Massachusetts Institute of Technology between September and October 2001. Fourteen male and female graduate students performed the experiment. The average age was 25, with a range between 22 to 31 years. There were three civilian pilots, holding a private pilot's license at a minimum, with an average of 456 flight hours. Subjects described their computer experience as "general" at least, with four subjects claiming to have "extensive" computer experience. Only one subject had experiences with flight management systems of both Boeing and Airbus aircraft. Subjects needed on average 3.2 hours to complete the experiment; approximately two of those hours were dedicated to completing a training tutorial, and another hour for running the data collection scenarios and for completing the questionnaires.

4.2 Route Cost Discussion

The route cost provided the primary quantitative and objective measure of human performance for the in-flight replanning task. Other quantitative measurements supported and added to the route cost findings, but did not stand-alone as a measurement of human performance like route cost. The actual route cost data used for the analyses was a transformed version of the raw route cost obtained as data output from the experiment. We transformed the raw route costs to help reduce subject performance variances and to better fit the data to a normal distribution.

For each subject, we first normalized the raw route costs against the associated minimal cost route for each scenario. As previously described in Section 3.5 (Independent Variables), subjects achieved their minimal cost routes from the scenarios without time pressures. Normalizing route costs helped to reduce the skewing effects from extreme subject performances. Having a normalized route cost equal to one represented achieving the subject's optimal performance in minimizing the route cost for a specific scenario. Furthermore, normalization of the raw route cost provided a reference for optimal performance, which allowed us to more naturally compare and discuss the quantitative findings of subject performance.

The normalized route cost had an approximately lognormal distribution; we needed a normal distribution for analysis. Therefore, we transformed the normalized cost by its natural logarithm to better fit a normal distribution, shown in Figure 4-1. A subject achieved the minimal cost route with a route cost equal to zero, or $\ln(1) = 0$; however, a route cost equal to zero did not indicate a cost-free route. For example, a route exposed to hazards, deviated from the TOT goal, or a combination thereof, could still have zero cost if that route was the minimal cost route. A route cost below zero indicated that under time pressure a subject achieved a lower cost route than in the case without any time pressure. While it was possible to have a route cost below zero, this did not occur often. On average, route costs under time pressures were significantly worse than without time pressure.

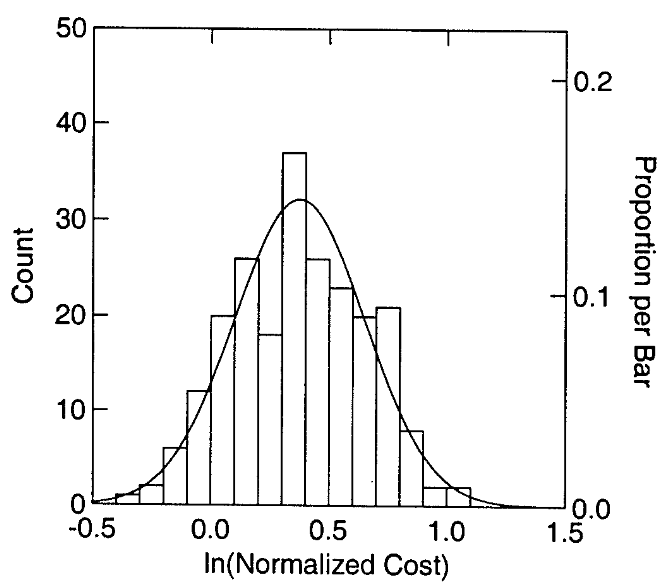


Figure 4-1. Approximate Normal Distribution of Transformed Data.

For brevity, the rest of this document refers to the normalized and log-transformed raw route cost as simply "route cost." The document specifically labels any other forms of route cost, such as the raw route cost or the normalized route cost.

As done in Section 3.6 (Dependent Variables), let us relate route costs due to actual hazards and TOT goal deviations to flight time equivalents. Again, we use a military flight scenario example, assuming a map display of 200 square nm, and a typical combat velocity of 800 nm per hour. Table 4-1 summarizes the important flight time equivalents of red and orange threats, using a least-squares linear regression. These flight time equivalents are only approximate, however, because route cost was a normalized value of the raw route cost. With the above assumptions, a route cost equal to zero was equivalent to 6.7 minutes in only red threats or 22.4 minutes in only orange threats. This corresponded to an increase of 1.0 minute in red threats or 3.2 minutes in orange threats for each 0.1 increase in route cost. The maximum acceptable deviation from the TOT goal was below the minimal route cost of zero average.

Table 4-1. Route Cost Flight Time Equivalent Summary.

	Route Cost = 0.0 Flight Time Equivalent (min)	Flight Time (min) per 0.1 Route Cost
Red Threat	6.7	1.0
Orange Threat	22.4	3.2
Yellow Threat	67.1	9.5

Figure 4-2 plots the flight time equivalent relationships for each threat level, with route cost on the horizontal axis and flight time on the vertical axis. For example, a route cost of 0.2 had a flight equivalent of 8.6 minutes through red threats only. Alternatively, a route cost of 0.2 had a flight equivalent of 28.7 or 86.1 minutes through only orange or yellow threats, respectively. Time in minutes relates route costs to tangible factors within the flight environment, which aids in the discussion of results and in understanding the severity of route cost differences. There are many real-world examples that can apply to the relatively abstract experiment, the in-flight replanning environment being only one.

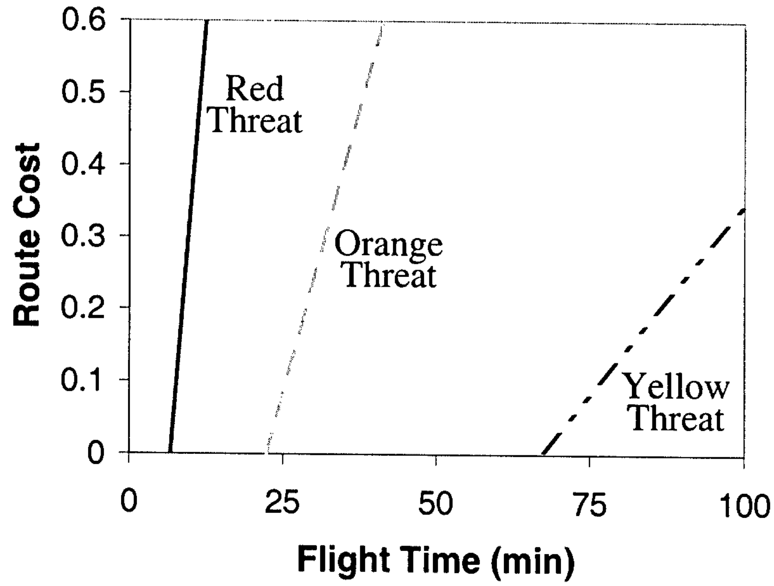


Figure 4-2. Route Cost Flight Equivalent Relationships.

4.3 Statistical Analysis Overview

We carried out a repeated measures analysis using a mixed regression and a two-way repeated measures analysis of variance (ANOVA). A mixed regression accommodated the Graeco-Latin Square test matrix design used to analyze the within-subjects main effects of time pressure, automation, and map complexity. By convention, results were significant if the test statistic gave at least a 95% confidence level in rejecting the appropriate null hypothesis, having a two-tail probability less than 0.05.

In general, a mixed regression tests the statistical significance of a best-fit line through a main factor having a slope equal to zero. A repeated measures ANOVA allows for statistical contrasts between specific test conditions. A repeated measures ANOVA uses a general linear model approach, similar to a mixed regression analysis. In general, an ANOVA evaluates the null hypothesis that the means of the groups are equal ($\mu_i = \mu_j = \ldots = \mu_n$) by examining the ratio of between-group variance (MS_{effect}) to within-group variance (MS_{error}) against the F-ratio distribution. When appropriate, the probability values were Huynh-Feldt corrected for failures in the assumption of general sphericity.

A mixed regression and repeated measures ANOVA must meet several similar assumptions for the results to be valid. First, we must assume the effects are additive and

constant, which means that a measured value is the sum of pieces. Figure 4-3 is a block diagram illustrating the additive model adopted to describe the factors influencing the experimental dependent variable. The route cost measurement, for example, was assumed to be dependent on the linear combination of subject, automation, time pressure, and map factors. The additive model implied that the main effects were independent from and did not interact with the subject effects. Because route cost was a measurement of a multivariate and complex task, the additive and constant effect assumptions were suspect [Cobb, 1998 (Chapter 12)]. However, taking the log-transform of the raw route cost most directly approximated an additive and constant relationship between route cost and the independent variables. The validation of the following assumptions indirectly supported the additivity assumption.

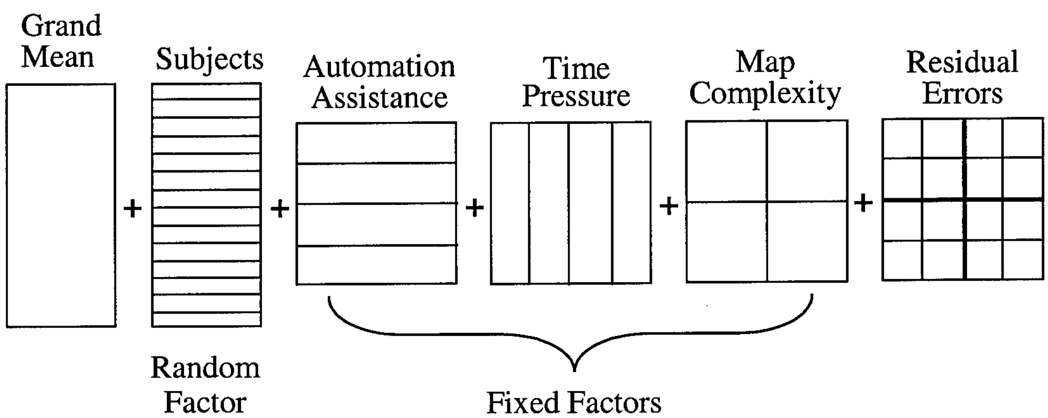


Figure 4-3. Additive Model for Dependent Variable (adapted from Cobb [1998]).

The analyses assume measurement errors (residuals) are independent and identically normally distributed. Within each main factor, we used normal probability plots to visualize the degree of normality in the respective distributions. Both route cost and route modification residual data followed nearly linear normal probability trends at each time pressure and automation category. The analyses also assume that standard deviations (SD) are equal within each group of a main factor. The ratio of maximum SD to minimum SD within a main factor should be less than three to meet this assumption of equivalent SD [Cobb, 1998 (Chapter 12)]; all relevant SD ratios in our analysis were less than three.

A mixed-interactions model applied to the route cost data set because it included both random and fixed factors. The analyses considered subjects a random effect, which meant that

subjects were sampled at random from the population of interest. Map complexity, automation assistance, and time pressure were fixed effects, effects seen only by the subjects in the experiment. There are at least two models used for data with mixed interactions, of which a repeated measures ANOVA follows the restricted model. As described by Cobb [1998 (Chapter 13)], the restricted mixed model assumes that:

1. for each random factor level (subject), the interaction terms add to zero over the level of each fixed factor—map complexity, automation assistance, and time pressure,
2. subject effects equal zero, each subject's mean is subtracted from his or her scores, and
3. the subject averages must be equal to zero

In addition, an ANOVA allows for easy statistical comparisons, or contrasts within our complex experimental design. Contrasts allow statistical comparisons between specific conditions of the independent variables not directly tested from an ANOVA or mixed regression. Because we only wanted to contrast specific time pressure or automation conditions, we needed to adjust the data for map effects to best approximate the effects due only to time pressure and automation. The additivity model allowed us to simply subtract out the map effects, using the overall map averages as the best estimator of map effects. We first calculated the overall averages for the four maps and their differences from the overall mean. Within each subject, we then subtracted out the difference from corresponding maps, adjusting each data point for map effects. With the map-adjusted data, the means of each map were now equal.

We used SYSTAT version 10 (SYSTAT Software, Inc.), statistical software package for all statistical data analyses, including the analyses of subjective data.

4.4 Within-Subject Main Effects

A mixed regression tested the within-subject main effects on route cost due to time pressure, automation assistance, and map complexity. We used contrasts to test for significant differences between selected groups within the main effects. Figures 4-4 and 4-5 show subject replanning performance grouped by automation category and time pressure. The route cost is on the vertical axis, where decreasing cost indicates increasing subject performance and a cost of

zero represents the averaged minimum route cost (not a zero-cost route). The horizontal axis categorically plots the automation assistance and time pressures. The error bars represent the standard error of the mean.

Figure 4-4 shows effects due to varying automation assistance, with standard error of the mean (0.021 average) indicated by the error bars. Overall, automation assistance had a significant effect on route cost, $z = 5.626$, $p < 0.0005$. Performance with Full was significantly the best, $F(1,13) = 187.3$, $p < 0.0005$. Route cost was three times less with Full (0.123) than with the average of None and Partial (0.459), equivalent to 3.4 minutes of flight time through only red threats. Performance with Hazard and None was significantly different, having a route cost 0.062 better with Hazard, $F(1,13) = 5.198$, $p = 0.040$. Only route cost with Constraint did not have significant differences from with None; the route cost was actually slightly higher with Constraint.

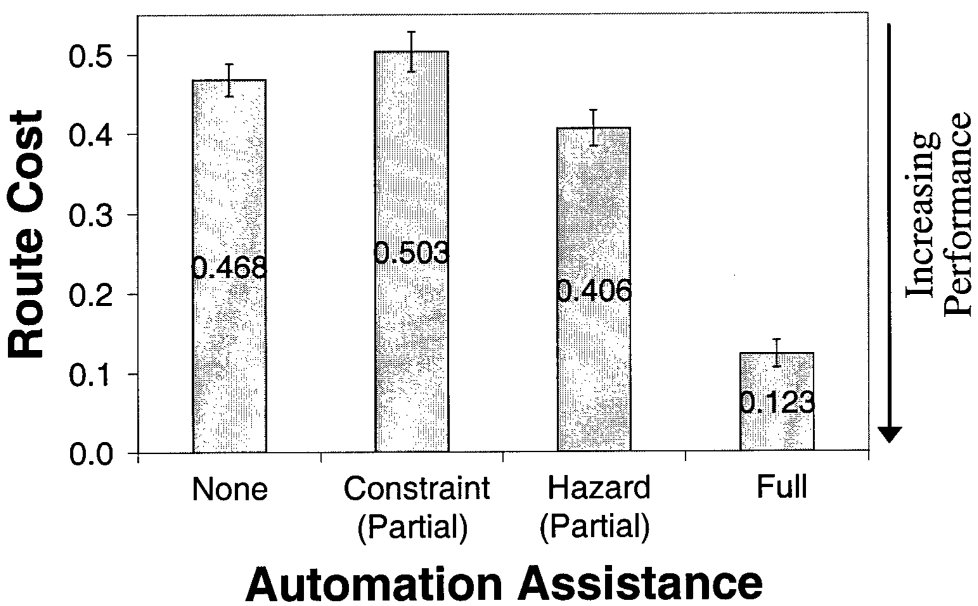


Figure 4-4. Automation Assistance Main Effects.

Figure 4-5 shows route cost effects due to varying time pressures, with standard error of the mean (0.029 average) indicated by the error bars. Overall, time pressure had a significant effect on route cost, $z = 2.886$, $p = 0.004$. The difference in performance was significant between 20 and 28 seconds, $F(1,13) = 8.764$, $p = 0.011$, decreasing in route cost by 0.089. Performance differences among 28, 40, and 55-second time pressures were not significant. The

average subject performance was the best at 55 seconds, with a route cost significantly less than at 20 seconds, $F(1,13) = 9.336$, $p = 0.009$. At 55 seconds, however, subjects were still far from their optimal performance (route cost = 0), with a route cost mean of 0.336. Between time pressures, trends suggested a non-monotonic increase in replanning performance. This trend is discussed further in Section 6.2, Time-Critical Implications.

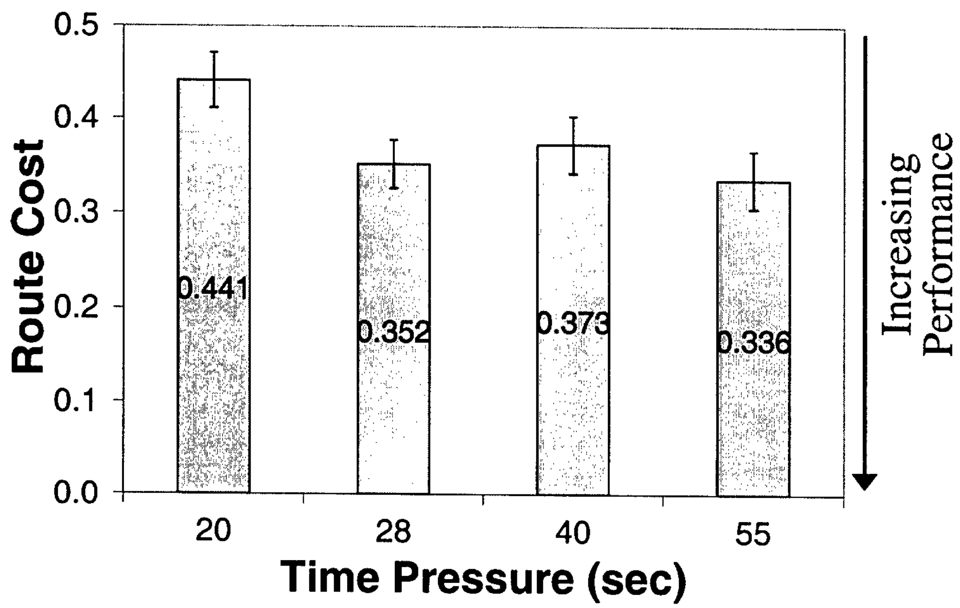


Figure 4-5. Time Pressure Main Effects.

There was a significant effect of map complexity on route costs, $z = 5.987$, $p < 0.0005$. Section 4.3 explains the route cost adjustments made to remove these map effects. In addition, learning effects from map familiarity through trials were a concern. We checked for learning effects by trial order within each map, and results showed that learning effects were not significant.

4.5 Within-Subject Interaction Effects

Of particular interest were the within-subject interaction effects between time pressure and automation assistance on route cost. As discussed previously in Section 4.3, Statistical Analysis Overview, we subtracted map effects out of the interactions to best approximate effects

solely due to time pressure, automation, and residual error. We used contrasts to test the significance of interaction effects.

Figure 4-6 shows the route costs at each combination of time pressure and automation assistance, averaged across all subjects, with standard error of the mean (0.025 average) indicated by the error bars. Lines connect data points from the same automation assistance category. The vertical axis represents the route cost, where decreasing cost represents increasing subject performance. A route cost equal to zero represents the averaged minimal cost route, or the subjects' average optimal performance. For example, the route cost at 40 seconds with Full (circles) is approximately 0.1, and with Constraint (diamonds) it is approximately 0.55. Table 4-2 summarizes quantitatively the route costs at each data point in Figure 4-6.

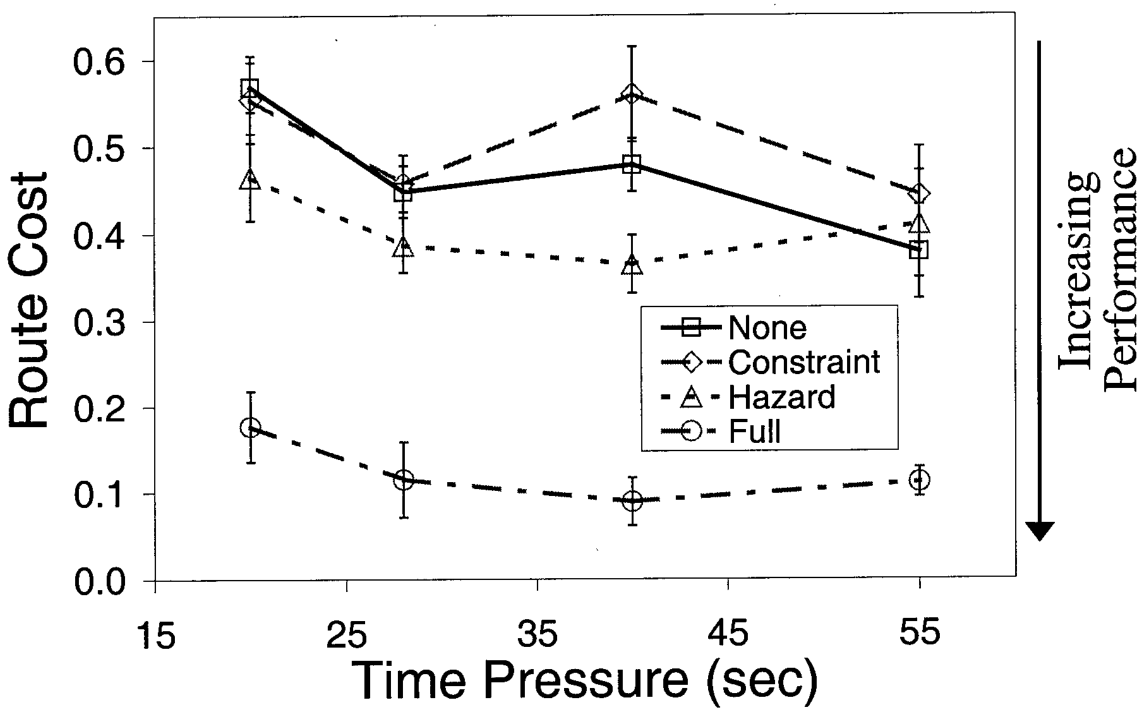


Figure 4-6. Time Pressure and Automation Category Interactions.

There were temporal performance interactions between Partial and None, while performance with Full was significantly the best at each time pressure, $F(1,13) = 18.30$, $p = 0.001$. Only with None did performance improve as more time was available, with significant route cost reduction from 20 seconds (0.568) to time scales greater than 20 seconds (0.435

average), F(1,13) = 6.25, $p < 0.03$. While not significant, performance trends with Partial and Full suggested that changes did occur in subject performance between time pressures.

Subjects performed better with Hazard than with Constraint and None at each time pressure except 55 seconds; however, this was only significant at 40 seconds, F(1,13) = 7.09, $p < 0.02$. This result supported the analysis of main effects in that performance with Hazard was better than with None or Constraint. Trends suggested that None had the highest route cost at 20 seconds, while it had a lower route cost than Partial at 55 seconds. At 20, 28, and 55 seconds, no significant differences existed between route costs with None or Partial.

Table 4-2. Route Cost Interactions Summary.

Route Cost	**Time Pressure (sec)**				
Automation	**20**	**28**	**40**	**55**	Grand Mean
None	0.568	0.448	0.478	0.378	0.468
Constraint	0.554	0.457	0.559	0.443	0.503
Hazard	0.465	0.387	0.364	0.410	0.406
Full	0.177	0.115	0.089	0.112	0.123
Grand Mean	0.441	0.352	0.373	0.336	0.375

4.6 Temporal Benefit from Automation Assistance

This section discusses the results supporting the concept of a limited temporal benefit from automation assistance. Recall from the Introduction chapter, well- designed automation should assist the human at extreme time pressures where there is not enough time for decision-making. As time pressure relaxes, the human should be better able to integrate the environment's diverse and complex information. Given enough time, for many applications an unaided human will produce a solution at least as good as when given automation assistance for similar tasks. We term the time needed to perform as good with None as with automated assistance the *characteristic time* (CT), or the time interval in which having automation assistance is beneficial. Characteristic time is highly dependent on the type and amount of automation assistance received by the subject.

Figure 4-7 shows the temporal benefit from automation assistance. The vertical axis represents automation benefit, and time is on the horizontal axis. For each time pressure, the automation benefit is the difference in route costs between a given automation level and None:

Automation Benefit = $Cost_{None} - Cost_{Auto\ Assist}$. Positive benefit indicates beneficial automation assistance, the more positive the better. Negative benefit indicates detrimental automation assistance, where subject performance was actually worse with automation assistance than with None. A benefit equal to zero represents no performance difference in having automation assistance. Assuming the benefit can be modeled with a monotonically decreasing function, the time at which zero benefit occurs indicates the CT, the time at which automation benefit crosses from positive to negative benefit.

Figure 4-7 shows the best-fit lines through each of the automation category data points. We used a least-squares linear regression to fit the data, and projected the line backwards to time equals zero. Overall, Full had the greatest benefit, $F(1,13) = 66.25$, $p < 0.0005$. The results suggested that Hazard provided the next greatest benefit over Constraint, while the difference was slightly insignificant, $F(1,13) = 4.456$, $p = 0.055$. Automation benefit results between time pressures did not have significant trends overall or within each automation assistance category.

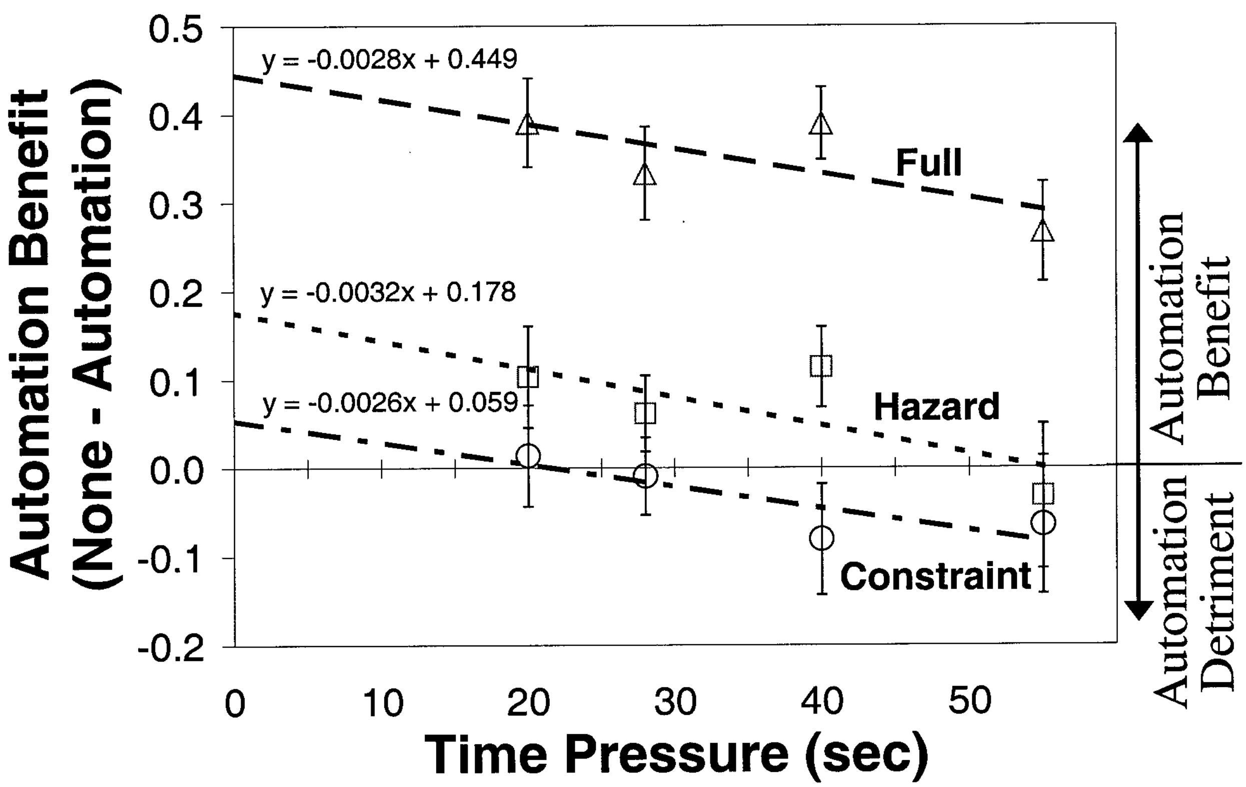


Figure 4-7. Temporal Benefit from Automation Assistance.

To provide a basis for description and discussion of automation benefit, we use a linear metric. However, the linear slopes are only slightly negative by observation; by a mixed regression analysis, the slight negative linear trends are not significantly different from a zero-slope line. The linear metric is simple and easy to understand, and fits the data well as seen by similar negative slopes for each automation assistance category.

While the linear metric models the time pressures between 20 and 55 seconds, we can only speculate outside the tested region. However, the vertical axis intercepts may describe the relative benefits received from automation assistance initially. For example, this intercept gives an indication of the automation benefit at the most extreme time pressures where there is no time for human inputs into the decision-making process. Referencing Figure 4-7, the backward projected trends suggest that any automation assistance provided at least some benefit over None at the extreme time pressures, with all positive automation benefits at zero seconds.

The actual initial route costs supported the suggested trends from automation assistance at the extreme time pressures. Table 4-3 summarizes the actual initial, or zero-second averaged route cost benefits from automation assistance. To calculate the initial route costs for each automation assistance category, we averaged across the four maps without regard to route acceptability, a process consistent with the route cost analysis. The initial route cost with None was 1.253, which was at best twice as costly as with Constraint, and was at worst six times as costly as with Hazard. While the zero-second benefit from using Full was slightly lower than with Hazard, Full suggested an initial acceptable route. The initial benefit with Hazard was the most because it avoided the high cost threat severity levels without regard to the TOT and fuel constraints.

Subjects did not reach the CT with Full within the tested time pressures. This meant that subjects did not perform as well unassisted as with having Full, even at 55 seconds. Assuming a linear trend continued, CT with Full would be 160 seconds. With Hazard, the CT was approximately 55 seconds, while with Constraint the CT was approximately 20 seconds. Subjects performed as well with as without Constraint at time pressures less than 20 seconds.

Table 4-3. Zero-Second Automation Benefit and Violations.

	Constraint	Hazard	Full
Route Cost Benefit	0.65	1.05	0.92
Constraint Violation	Yes	Yes	No

4.7 Analysis of Mission Failures

An analysis of mission failures provides a gross look at subject performance: the mission either succeeded or failed. A mission failed if at least one of the following occurred: the route intercepted the highest-level hazard (brown), the subject arrived at the target outside the acceptable time window, or the planned route did not have enough fuel at the egress point. Overall, there was a 14.3% mission failure rate, or 32 of the 224 missions attempted. While the analysis of mission failure may be highly dependent on a few subjects, we believe these results are not. Figure 4-8 is a histogram of scenario failures per subject. The average number of mission failures per subject is slightly over two per subject, with one subject failing six and another failing zero at the extremes.

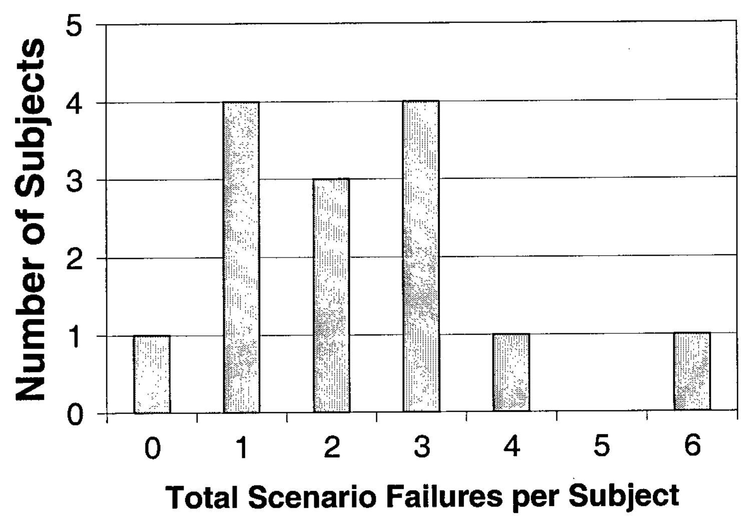


Figure 4-8. Mission Failure Histogram.

Table 4-4 summarizes the mission failure analysis. Most mission failures occurred with Hazard (13 failures), which was contrary to the results from main effects. There were five failures with Full, each induced solely by the subject since Full suggested an acceptable initial route. A different perspective on mission failures was that all initial routes suggested by Partial and None were unacceptable, which was 56 trials each. Therefore, subjects with Hazard were able to fix approximately 75 percent of the trials with initially unacceptable routes. Similarly, subjects made 90 percent and 85 percent of the None and Constraint route suggestions acceptable. Subjects had the fewest mission failures given 40 seconds (4 failures), and the most

failures at 55 seconds (11 failures). Contrary to the route cost analysis trends, the number of failures at 20 and 55 seconds were approximately equal. It was difficult to interpret the number of mission failures more specifically, within each time pressure or automation assistance category, because the results were too dependent on individual maps.

Table 4-4. Mission Failure Count Summary.

Mission Failure Count	Time Pressure (sec)				
Automation	20	28	40	55	Grand Total
None	2	1	1	2	6
Constraint	0	3	2	3	8
Hazard	6	2	0	5	13
Full	2	1	1	1	5
Grand Total	10	7	4	11	32

Table 4-5 summarizes the constraint violation frequency with respect to each time pressure and automation type. There were more constraint violations than mission failures, as a single mission failure could include one or more constraint violations. In total, there was approximately the same number of each constraint violation, ranging between 11 and 15 violations. At 20 seconds, we observed the most fuel violations (8 violations), approximately four times as many as at other time pressures. Brown hazard and TOT violations were most frequent at 55 seconds. Subjects violated each constraint the least at 40 seconds.

By observation, constraint violation trends suggested that mission failures related to the information elements not integrated by the route automation assistance. With Constraint, subjects mostly violated the brown hazard constraint; only one of the eight violations was not a brown hazard violation. Subjects with Hazard violated fuel and TOT constraints the most. Lastly, there were only fuel constraint violations with Full, and None had similar numbers of each violation.

Table 4-5. Constraint Violations Summary.

Violations	Time Pressure (sec)				Automation Assistance				Total
	20	28	40	55	None	Constraint	Hazard	Full	
Brown Hazard	3	3	1	6	2	7	4	0	13
Fuel	8	2	2	3	2	0	8	5	15
TOT	3	2	1	5	3	1	7	0	11

4.8 Route Modification Events

The analysis of route modification events gave a perspective into the human physical and cognitive interactions with automation and time pressure. A route modification included any route movements that changed the route trajectory. Simply adding a waypoint inline and along the route was not a route modification. This analysis augmented the quantitative route cost and subjective data analyses; used alone, however, it was not a complete measure of human performance or cognitive workload in the replanning task. As described in Section 4.3, we adjusted the data to remove map complexity effects, and used a repeated measures analysis.

Figure 4-9 shows the number of route modification events at each automation and time pressure combination, averaged across all subjects, with an averaged standard error of the mean of 1.2 modifications. A line connects the data points within the same automation assistance category. The vertical axis represents the number of route modification events, and time pressure is on the horizontal axis. At 40 seconds, for example, there were 10.5 modifications on average with Hazard (triangles) and 4.9 modifications with Full (squares). Table 4-6 summarizes quantitatively the average route modifications per second, an inverse perspective from what Figure 4-9 shows.

Overall, automation assistance and time pressure had significant effects on the number of route modifications, $z = 5.564$, $p < 0.0005$ and $z = -12.23$, $p < 0.0005$ respectively. Full had significantly the least modification events (21.9), $F(1,13) = 63.674$, $p < 0.0005$, which directly translated into Full having the lowest overall modification rate (0.15 per second). With each increase in time increment, modifications significantly increased, $F(1,13) = 13.102$, $p < 0.003$. This corresponded to a nearly constant route modification rate among time pressures, ranging between 0.21 and 0.24 modifications per second, with a 0.22 per second overall average.

At 20 seconds, there were no significant differences among automation types, with the baseline 4.6 modifications average. Route modifications with None and Partial continued to show nearly identical increasing trends at 28 and 40 seconds; while route modifications with Full were significantly the least at 28 and 40 seconds, $F(1,13) = 49.094$, $p < 0.0005$. At 55 seconds, route modifications with None (11) and Full (9.9) were significantly lower than with Partial (15.7 average), $F(1,13) = 24.258$, $p < 0.0005$.

Of particular interest from Table 4-6 are the modification rate trends observed with Full. With Full, the number of route modifications was similarly low from 20 to 40 seconds (4 average), and then more than doubled at 55 seconds (9.9). This modification trend caused the initial significant decrease in modification rate from 20 seconds (0.22 per second) to 28 seconds (0.10 per second), $F(1,13) = 45.513$, $p < 0.0005$. We then observe a significant rate increase with Full from 40 seconds (0.12 per second) to 55 seconds (0.18 per second), $F(1,13) = 11.805$, $p = 0.004$. In addition, at 28 and 40 seconds, the rate with Full (0.11 per second average) was less than half the rate with None or Partial (0.25 per second average), $F(1,13) = 49.094$, $p < 0.0005$.

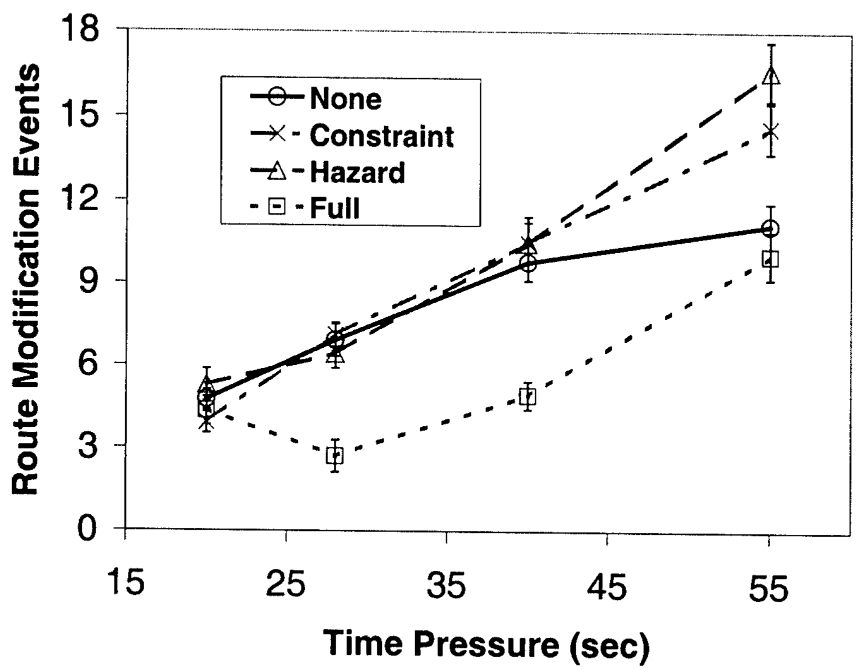


Figure 4-9. Route Modification Events versus Time.

Table 4-6. Route Modifications Per Second Summary.

Modifications/Sec	**Time Pressure (sec)**				
Automation	**20**	**28**	**40**	**55**	Grand Mean
None	0.24	0.25	0.24	0.20	0.23
Constraint	0.20	0.25	0.26	0.27	0.24
Hazard	0.26	0.23	0.26	0.30	0.26
Full	0.22	0.10	0.12	0.18	0.15
Grand Mean	0.23	0.21	0.22	0.24	0.22

5. SUBJECTIVE DATA

The subjective data provided an invaluable insight into human perceptions concerning replanning performance, automation assistance, and the relationships between the flight environment information elements. Analyzing the subjective data against the automation assistance types and time pressures gave a better understanding of the quantitative results. All the subjective data came from the experimental questionnaire responses using a discrete scale with possible values between one and five.

5.1 Statistical Analysis Overview

The nonparametric Friedman and sign tests were used to analyze the subjective data. The Friedman test determined if there were rank differences in performance between the levels of a given independent factor, such as automation or time pressure. We used the Sign test to specifically analyze pairs of levels within a factor, such as the None and Hazard pair. Parametric tests would not be appropriate for this ordinal data set because the responses were limited to five discrete ranks, in which many rank ties occurred. Since the subjective data did not follow a normal distribution, the more lenient assumptions underlying nonparametric tests matched our experimental results better.

The Friedman parametric equivalent is a two-way analysis of variance, as is the t-test for dependent samples for the sign test. By convention, nonparametric tests typically use the median rather than the mean because it is a more robust estimate than the average when there are skewed distributions or outliers [StatSoft, 2002]. The Friedman test statistic is approximately Chi-Square (χ^2) distributed. For reference, some of the Chi-Square statistical milestones that apply are: $\chi^2 = \{7.8, 9.4, 11.5, 16\} \approx p < \{0.05, 0.025, 0.01, 0.001\}$.

We used SYSTAT10 to create box plots that show the relationship between subject responses and the independent variables (time pressure and automation assistance). Box plots give a quick and comprehensive summary of the data [Moore, 1997]. The central box depicts the interquartile range, or 25% to 75% of the data. Lines extend beyond the central box to show the smallest and largest observations not considered outliers. A hinge within the central box depicts

the data median. Lastly, the box plots show suspected outliers as individual points, represented as circles and stars in the following box plots.

5.2 Question 1: Replanning Performance

To what degree were you able to replan within the time pressure?

1	2	3	4	5
Completed an optimal route	Completed a near optimal route, minimal improvement needed	Completed a good route, some improvement needed	Significant route improvement needed	Unable to improve or complete route

The purpose of the replanning performance question was to capture the subject's perception of their replanning performance. We analyzed the responses against automation assistance categories, time pressures, and actual performance. Figure 5-1 shows the box plots for subject responses to Question 1 versus automation assistance categories and time pressures. On the vertical axis are actual subject responses, where a response of one indicated the perception of developing the optimal route. The horizontal axis categorically labels the independent variables.

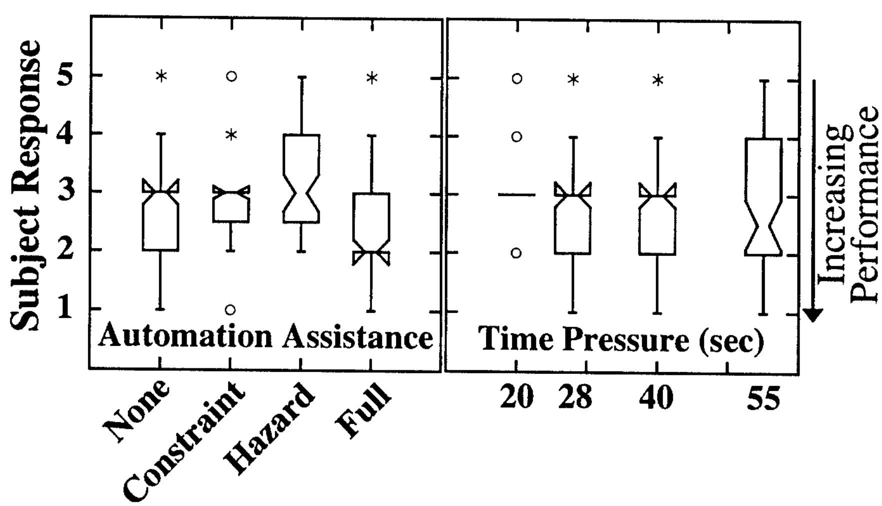


Figure 5-1. Replanning Performance Question Box Plots.

Perceived performance with None was not significantly different than with Partial or Full. Subjects perceived performance with Full significantly better than with Partial, $p < 0.002$. Route cost trends suggested that perceived performance was best at 55 seconds, while the only

significant difference was between 20 and 40 seconds, $p < 0.021$. Lastly, there was greater disagreement in perceived performance as time increased.

5.3 Question 2: Automation Assistance

To what degree did the suggested route assist you in the replanning task?

1	2	3	4	5
Relied completely on suggested route	Suggested route provided great assistance	Suggested route provided some assistance	Suggested route provided very little assistance	Suggested route provided no assistance

We wanted to capture the subject's perception of having automation assistance in the form of a suggested route. We also analyzed this question against automation assistance categories, time pressures, and actual performance. Figure 5-2 shows the box plots for subject responses versus automation assistance categories and time pressures. Subjects perceived Full to be significantly the most helpful, $p < 0.0005$. Responses equal to one or two defined the interquartile range for Full, indicating from great assistance to complete reliance. The median was three for None and Partial, indicating "some" assistance. There were no significant differences in perceived automation assistance between time pressures, each with a median of three and with similar response variances.

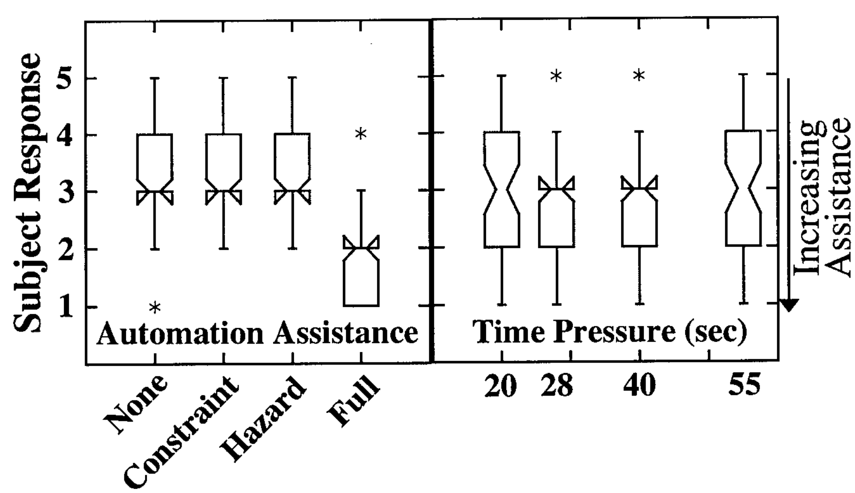


Figure 5-2. Automation Assistance Question Box Plots.

5.4 Correlations with Actual Performance

Figure 5-3 shows perceived performance and automation assistance correlations with actual performance. For the replanning performance question, actual performance positively correlated with perceived performance, Pearson correlation coefficient equal to 0.369. The correlation was strongest for subjects giving a score between two and three, suggesting subjects did not accurately perceive performance at the extreme scores. For the automation assistance question, actual performance positively correlated with perceived automation assistance. There was a significant linear correlation for both questions; however, the automation assistance question had greater positive correlation strength, with a Pearson correlation coefficient equal to 0.516.

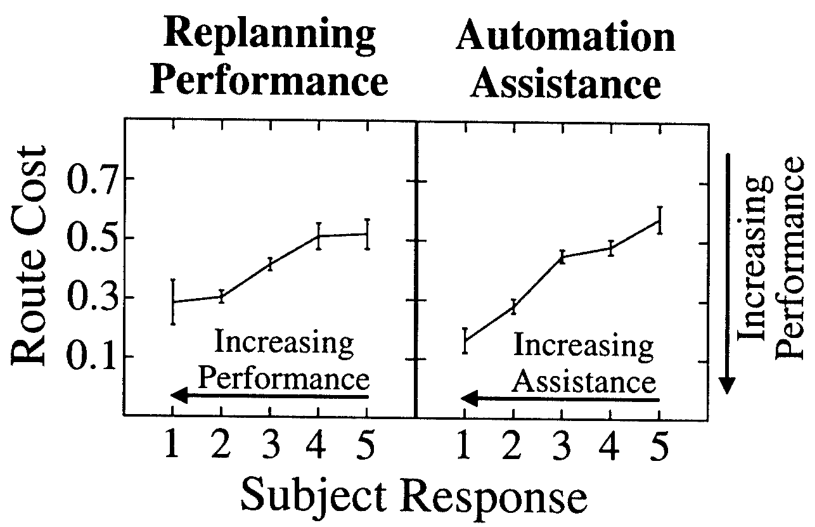


Figure 5-3. Actual Performance Correlations with Q1 & Q2 Responses.

5.5 Question 3: Information Element Difficulty

To what degree did each element restrict your ability to develop an optimal flight plan? The information elements were *time* pressure, *hazards*, *fuel* constraint, and *TOT*.

1	2	3	4	5
Not at all	Slightly	Somewhat	Significantly	Completely

In this section, we labeled the information elements as *time*, *hazard*, *fuel*, and *TOT*. There are several different ways to analyze this question, by information element or by time

pressure and automation. Figures 5-4 and 5-5 show box plots comparing an information element between groups of time pressure or automation. Subject responses are on the vertical axis, where increasing numbers indicate being more restrictive to developing an optimal flight plan. The horizontal axis categorically plots the automation assistance and time pressures. Significant results indicate a difference in perception of relative rankings, or relative restrictiveness.

Between automation assistance categories (Figure 5-4), there were significant differences in subject responses for hazard and fuel information elements, χ^2 (3) = 8.25, p < 0.041. *Time* and *TOT* information did not have significant differences between automation types, in fact having constant medians of four and three respectively. *Hazards* were most restrictive when replanning with Constraint, p < 0.031. *Fuel* was significantly more restricting with Hazard than with None and Constraint, p < 0.006.

Between time pressures (Figure 5-5), there were significant differences in subject responses for *time* and *TOT* information elements, χ^2 (3) = 9.150, p < 0.027. Subjects perceived the 20 and 28-seconds time pressures as significantly more restrictive than the 40 and 55-seconds time pressures, p < 0.004. There were no significant differences in perceived relative restrictions of *hazard* or *fuel* across time pressures, with a nearly constant median of three for both.

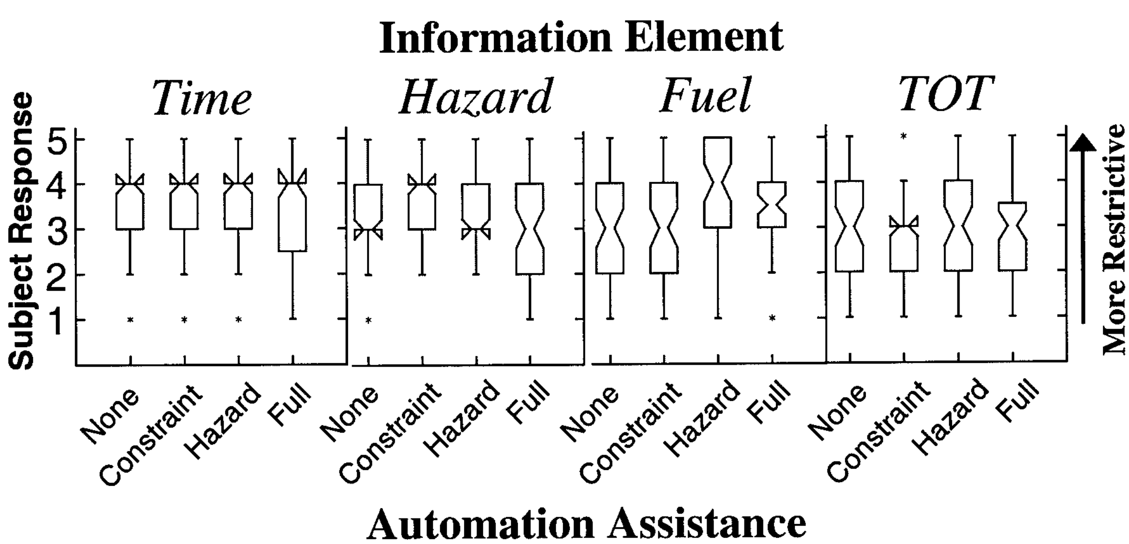


Figure 5-4. Information Elements Box Plots between Automation Categories.

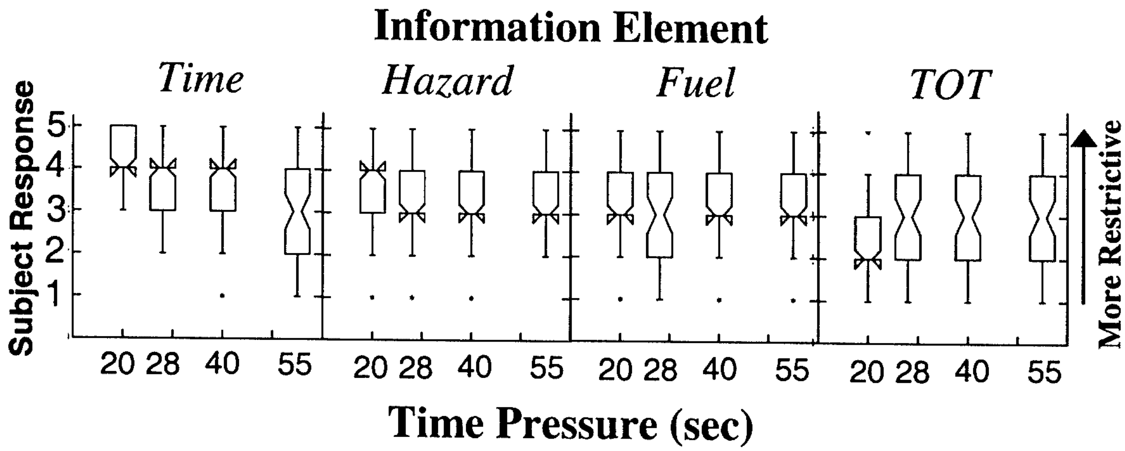


Figure 5-5. Information Elements Box Plots between Time Pressures.

Figures 5-6 and 5-7 show box plots comparing information elements within each group of time pressure and automation, respectively. Within each group, the box plots show from left to right: *time*, *hazard*, *TOT*, and *fuel* information elements. The vertical axis shows subject response, where increasing scores indicate being more restrictive. The horizontal axis categorically labels each group of time pressure and automation.

Within automation assistance categories (Figure 5-6), overall perceptions of relative restrictiveness of information elements were significantly different with Constraint and Hazard, χ^2 (3) = 16.007, $p < 0.001$. With None and Full, subjects were not able to perceive a significant difference between the information elements. With Constraint, perceptions of *TOT* and *fuel* were significantly the least restrictive, $p < 0.039$. With Hazard, trends suggested that *hazard* was less restrictive than *time* and *fuel*, while not significant.

Within time pressures (Figure 5-7), there were only significant differences in relative rankings between information elements at 20 and at 28 seconds, χ^2 (3) = 9.471, $p < 0.024$. At 20 seconds, *time* pressure was significantly more restricting than *fuel* or *TOT*, and *TOT* was significantly the least restricting, $p < 0.001$ and $p < 0.012$ respectively. At 28 seconds, *time* was again more restrictive than *fuel* and *TOT*, $p < 0.021$. Lastly, all information elements had the same subject response median equal to three at 55 seconds.

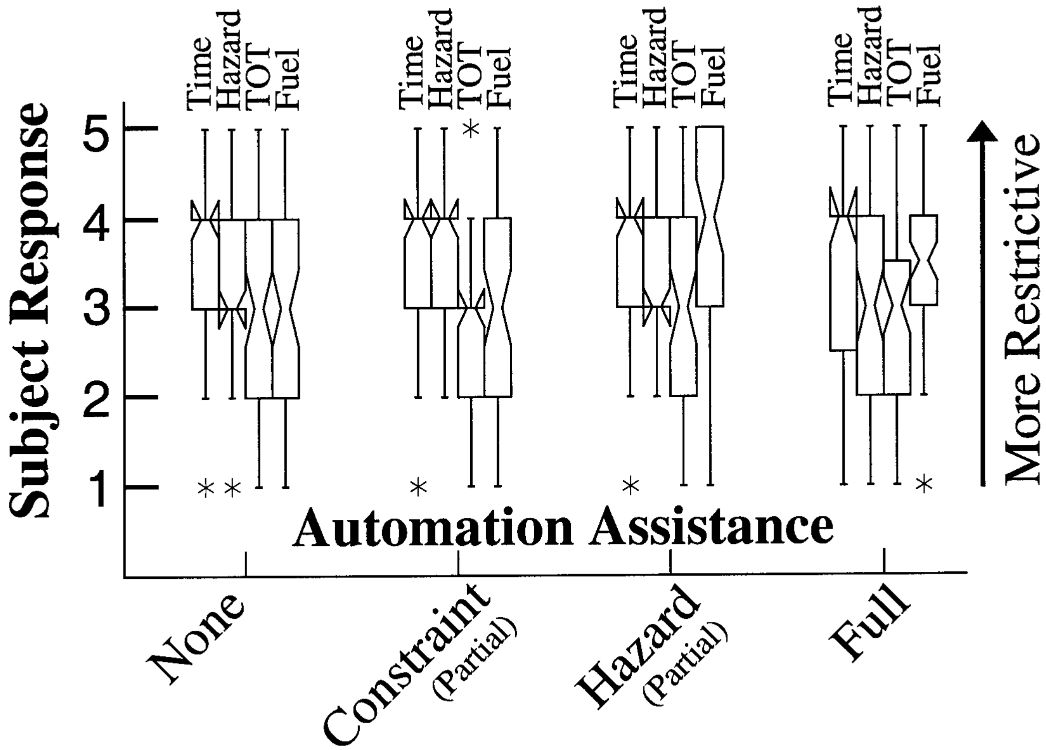


Figure 5-6. Information Elements Box Plots by Automation Categories.

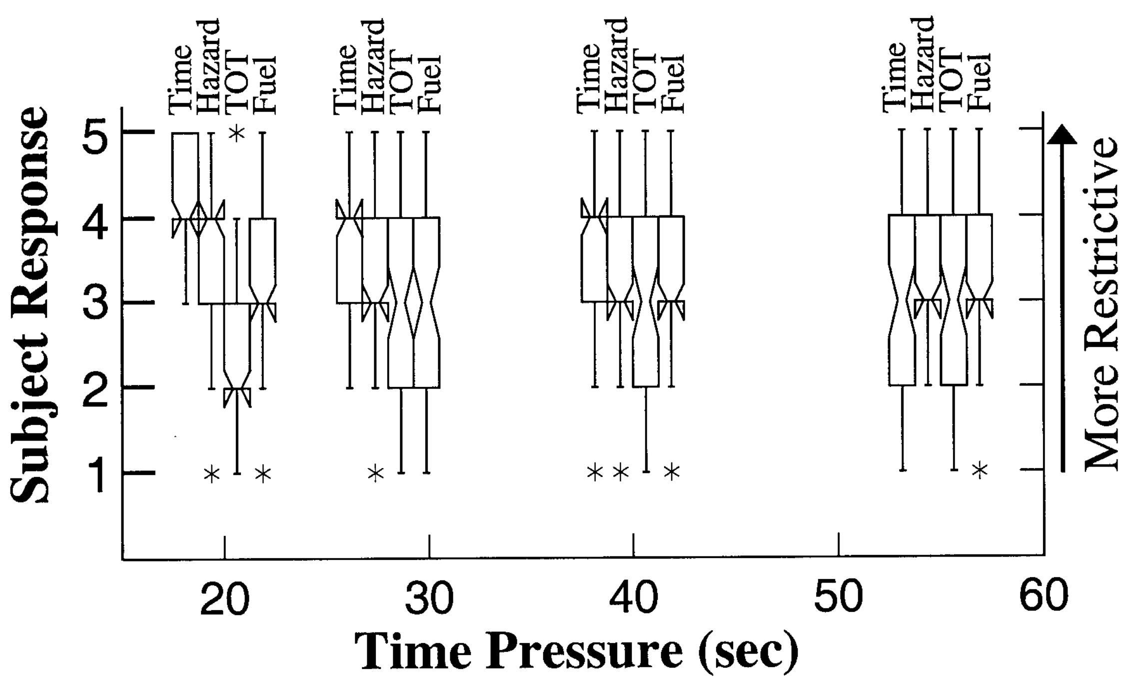


Figure 5-7. Information Elements Box Plots by Time Pressures.

5.6 Qualitative Data (Subject Comments)

This section will highlight the qualitative findings from a post-experiment short interview and questionnaire. The experimental proctor used the post-experiment questionnaire as a guide for a quick discussion on the experiment, and manually recorded subject responses. The questions asked subjects to make observations or comments about the adequacy of training, the replanning task, the route replanning strategies, the route automation assistance categories, and other general comments related to the experiment. A complete and detailed transcription of subject responses to each question can be found in Appendix D.

In response to the training adequacy question, subjects on average felt the training tutorial "mostly" prepared them for the experiment, giving a four out of five rating. When asked on how to improve the training, six subjects agreed that the training needed more time for practice. Four subjects specifically wanted more practice scenarios included in the training. While many subjects wanted more training, given that there was a finite amount of time for the experiment, they agreed the training adequately prepared them for the problem difficulty encountered in the data collection scenarios. Furthermore, subjects were anxious to begin the actual simulation after two hours of training; more practice may not have been an option.

There were several common responses observed in the qualitative analysis. In general, subjects used the route replanning strategies and goals learned from the training session. The strategies were mostly dependent on the time available and the given automation assistance, with a basic strategy to satisfy the mission constraints first. Twelve of the fourteen subjects cited preferring replanning with Full the most; the other two subjects did not specifically comment. With Full, several subjects further stated they simply made small route adjustments where they felt improvement was possible. In addition, subjects overwhelmingly used the suggested route as the initial starting point for replanning. Only two subjects used the revert function to recall the original route, one time each.

We asked all subjects to give their automation assistance preferences, ranking from most to least liked if possible. As just mentioned, subjects preferred replanning with Full the most. The greatest disparities were in subject responses to preferences between replanning with Partial and None. Table 5-1 summarizes subject preferences between Partial and None automation assistance only. For example, eight subjects preferred replanning with Constraint to replanning

with Hazard or None. Eleven subjects preferred having Partial to None. Constraint appeared to be highly preferred to either Hazard or None. Commenting on why they preferred Constraint to Hazard, five subjects claimed that hazard avoidance was easier than trying to meet fuel and TOT constraints. Similar number of subjects preferred None and Hazard the least, five and four subjects respectively. In support of route automation assistance, four subjects claimed that it gave a predictable perspective on the impending problems.

Table 5-1. Partial and None Subject Preference Summary.

	None	Constraint	Hazard
More Preferred Count	2	8	3
Least Preferred Count	5	0	4

A few comments were insightful into the overall experiment. One subject said, "Having more time got me focused on cost optimization, and forgetting about goals and that [route] optimization may change [the mission] constraints." Another subject concurred. In addition, one subject commented on feeling they made the suggested route worse at times when given Full automation assistance. Finally, one subject stated that while automation assistance reduced uncertainty, it did not necessarily reduce the replanning workload.

[Page Intentionally Blank]

6. DISCUSSION

Chapter 6 discusses and develops the important relationships and ideas from the objective and subjective results. This chapter will highlight the statistically significant results, and some of the most important or unexpected findings that may be inferred from trends that were not statistically significant. When appropriate, we will generalize these findings and ideas to applications other than in-flight replanning. In addition to route cost, the analysis of mission failures and subjective data provided a valuable perspective on subject replanning performance, and differed from the route cost trends in some conditions. In general, subjective findings supported the objective findings regarding replanning performance, automation assistance preferences, and information element interactions.

6.1 Automation and Time Pressure Results

For automation assistance, objective and subjective results with Full automation assistance had the most consistent and conclusive findings. Overall, actual replanning performance with Full was significantly the best and provided the most benefit over None. Furthermore, subjects did perceive performance with Full better than with Partial. Mission failures with Full were the least. Subjects significantly ranked Full the most assisting automation, providing between "great assistance" and "relied completely" upon the automation. Lastly, subjects overwhelmingly stated preferring replanning with Full the most.

Full combined the automation assistance capabilities of both Hazard and Constraint. Interestingly, the automation benefit from Full was greater than the sum of its individual automation modules, the sum of Hazard and Constraint benefits. With the linear automation benefit versus time trends, Full benefit was nearly double the summed Partial benefit at each time pressure, with zero-second route cost benefits over None equivalent to 0.449 and 0.237 respectively. This automation benefit difference between Full and Partial was equivalent to approximately two extra minutes of flight time through a red threat. This finding suggested a compounding benefit from decision-aiding automation that more fully integrated the processing of information.

At 20 seconds, any automation assistance was better than None, which indicated that automation assisted the human when time pressures most likely limited their cognitive ability to replan unaided. Interestingly, with Full and Partial, significant performance improvements did not occur among time pressures, even between 20 and 55 seconds. We observed similar trends in perceived replanning performance with no significant differences in rankings between time pressures. Only with None did performance significantly improve from 20 seconds to time scales greater than 20 seconds. Within the tested 55 seconds, performance with automation assistance did not change significantly with time; indicating that the time-critical limit for performance with automation was beyond 55 seconds for this problem difficulty.

The mission failure trend from time pressures of 20 to 40 seconds followed expectations, with decreasing failures from 10 (20 sec) to 7 (28 sec) to 4 (40 sec). Surprising, however, was the number of mission failures at 55 seconds (11 failures) being similar to that at 20 seconds (10 failures). Subject rankings perceived the 20, 28 seconds time pressures as significantly more restrictive to flight planning than the 40, 55 seconds time pressures: so why the failures at 55 seconds? With 55 seconds, subjects had time to significantly improve route costs. However, as qualitative findings suggested, with more time (55 seconds), subjects would sometimes forget about satisfying the mission constraints. With only a few seconds remaining, subjects would then realize the constraint violation, and most likely scrambled unsuccessfully to satisfy the forgotten route constraints. More time induced constraint violations in part because humans more focused attention on route optimization than on constraint violations.

The route cost and mission failures analyses did not always agree. The route cost analyses showed overall replanning performance with Hazard was significantly better than with None. However, subjects with Hazard failed the most missions, with double the number of failed missions than with None, 13 and 6 respectively. In addition, the least mission failures occurred at 40 seconds, with less than half the failures at 55 seconds; this was contrary to route cost trends. This inconsistency made it difficult to make conclusive remarks about replanning performance among time pressures, and between None and Partial. Combining with route cost the results from mission failures could provide a more complete and comprehensive metric for replanning performance.

In opposition to the proposition that route cost was a valid metric for replanning performance was the fact that route costs from failed missions were included in the route cost

analysis. Supporting the validity of the route cost was the extensive training subjects received in being motivated to satisfy mission constraints before minimizing route costs. Therefore, we assumed that in failing missions at higher time pressures (20, 28, 40 seconds), the subject did not have time to meet the constraints, let alone worry about the route cost. Thus we observed the highest route cost at 20 seconds. In addition, the failures at 55 seconds arguably were due to insignificant constraint violations, induced from minor errors in replanning optimization (i.e., just slightly skimming the edge of a severe threat). Route cost at 55 seconds reflected this mission failure trend because subjects were not able to significantly improve route cost from the 28 and 40-seconds time pressures, even though qualitative findings suggested there was adequate time to improve route cost at 55 seconds.

Mission failures were a gross measurement of performance, unable to capture the subtle performance trends as well as the route cost metric. Furthermore, with only a 14.3% failure rate, it was difficult to make conclusive statements about replanning performance based solely on a mission failure analysis. We were confident that route cost was the best objective measure used from the experiment to relate automation assistance and time pressures with replanning performance.

6.2 Time-Critical Implications

Between time pressures, trends suggested a non-monotonically increasing relationship with overall replanning performance. This trend demonstrated the idea that within a time-critical domain, humans may actually perform worse under certain time pressures than at higher time pressures. A likely explanation was that at some time pressure, subjects felt there was enough time to make major route modifications to the automated route suggestion in efforts to significantly reduce its route cost; but consequently were not successful. An example of a major route modification would be to take the route through threats much differently than what the automation initially suggested. Alternatively, with local route adjustments, subjects would slightly modify the initial route, minimizing the initial route's hazard exposures or TOT goal deviations.

Route automation assistance gave subjects a predictable starting point because they understood how the route automation assistance behaved, regardless of their automation

preferences. In efforts to find a better global solution by searching far from the initial route, subjects would be making risky and less informed decisions because they were without an accurate reference of route acceptability. Now, with major deviations from the initial route, subjects would then find themselves without enough time to find and complete an acceptable global solution, or to go back to the initial suggested route. Therefore, the final route could be worse than had the subject remained more conservative in modifying the suggested route. Enough time pressure would instead compel subjects to replan safely and conservatively, making only local and predictable route modifications to the initial route.

This idea, however, may only be relevant under time intervals within an initial time-critical period where subjects do not have enough time to reach optimal or near-optimal performance, such as 55 seconds in our experiment. Enough time to reach optimal performance would no longer force subjects to evaluate trade-offs between competing goals, negating the potential to misjudge their time-constrained replanning abilities. Route cost trends from 28 to 40 seconds, with a route cost increase from 0.352 to 0.372, suggested that 40 seconds was enough time to compel subjects to globally explore from the initially suggested route, but without success in reducing cost. In addition, we observed that mission failures at 55 seconds (11 failures) were more than double the mission failures at 40 seconds (4 failures).

The route modification rate with Full further supported the explanations for the non-monotonic increase in performance with time. Full had significantly the lowest modification rate average of 0.15 per second. The most likely explanation was that Full suggested an acceptable and very low cost route, a route in which subjects were hesitant to adjust. The subjective and qualitative findings also showed a reliance on the suggested route by Full. At 55 seconds, however, the route modification rate with Full (0.18 per second) significantly increased from the 28 and 40-seconds time pressures (0.11 per second average). Subjects most likely felt that 55 seconds was enough time to globally search for a better route than initially suggested by Full. Despite the route modification rate spike at 55 seconds, indicating obvious efforts to better this automated route suggestion, route cost with Full failed to significantly decrease from 20, 28, and 40 seconds to 55 seconds.

While not conclusive, the observed performance trends support a general description of the replanning strategies used within the tested time-critical scale, from 20 to 55 seconds. Among time pressures, 20 seconds most restricted overall performance, and trends suggested the

lowest route cost was at 55 seconds. Overall, the only significant performance improvement occurred between 20 and 28 seconds, dropping in route cost from 0.441 to 0.352. This indicated that approximately 28 seconds was the time needed to transition the replanning strategy from primarily satisfying mission constraints to both satisfying constraints and minimizing route costs. Between 28 and 40 seconds, we observed a reticence to stray dramatically away from the initial route while continually minimizing the route cost. Greater than 40 seconds, observations suggested another transition in replanning strategy from small to major route modifications in search for the global minimal cost route. We assume the major route modification tendency extends beyond the tested 55 seconds, and that the human will eventually lose the time-critical benefits from the initial automated route suggestion.

6.3 General Discussion

We expected to see the type of automation assistance influence what information was more restrictive in replanning the optimal route. Subject rankings from both the question on information element difficulty and the mission failure analysis gave a perspective into the relationship between automation assistance and information elements. Subjects significantly ranked replanning with Constraint the most restricted by the *hazard* information. Furthermore, trends strongly suggested that with Hazard, *fuel* information was most restrictive. In addition, most brown-hazard constraint violations occurred with Constraint, and the most fuel and TOT constraint violations were with Hazard. We observed through subjective rankings and mission failures that the route automation assistance did minimize the relative difficulty of the information elements it integrated.

With such a complex and time-constrained experiment, how do we know the experimental design was valid? The experimental results indirectly evaluated the integrity of the experimental design with respect to the chosen decision-making difficulty, time pressures, and automation assistance levels. With multiple experimental variables designed within the Graeco-Latin Square, having statistically significant results with only 14 subjects in both automation and time pressure factors attested to the integrity of the experimental design. The significant results from this experiment also provide a reference point for future research in this area of automation assistance to support time-critical decision-making.

The discrepancies between objective and perceived replanning performance highlight an important automation design issue: whether to design automation based on objective results or subject perceptions. Most subjects preferred Constraint to Hazard, yet overall route cost trends suggested that subjects performed better with Hazard than with Constraint. Subjects also least preferred replanning with Hazard and None equally; although, Hazard had an average route cost (0.406) significantly lower than None (0.468). For this discussion, we ignored the mission failure analysis because it neither supported nor contradicted subjective performance. Regarding inconsistencies between actual route cost and subjective performance, Sanders and McCormick [1993 (Chapter 10)] suggested designing using actual performance results, rather than using subjective preferences. While their argument was for an experiment looking at performance with arrangements between controls and displays, we also suggest using the objective route cost results.

The route modification count and rate gave a perspective into the relative difficulty and manual workload associated with route automation assistance. The overall route modification rates with None and Partial were similar (0.25 per second average), and were 0.1 per second more than with Full. Subjective results agreed that route assistance from Partial was the same as from None, and was not as assisting as Full. We know that None and Partial suggested an initially unacceptable route that had at least one constraint violation; and from qualitative data, we also know that subjects overwhelmingly began replanning with modifications on the initially suggested route. Therefore, the disparity in modification rate with None and Partial from Full may indicate the increased difficulty in replanning an initially unacceptable suggested route to make it acceptable.

As previously discussed, subjects commented on preferring Constraint to Hazard or None. Subjects most likely preferred Constraint to Hazard because with Constraint, the primary task of hazard exposure reduction was a perceptual task. With regard to visual modality for information processing, a study by Kirlik et al. [1996] found that perceptual augmentation of displays facilitated dynamic decision-making performance due to possible perceptual and pattern-recognitional heuristics. While with Hazard, subjects raced against the clock to remove the constraint violations inherent in the suggested route. This was primarily a mental workload task, which required a more complex estimate integrated along the whole route. These results,

supported by the Kirlik et al. study, suggested that subjects more easily processed perceptual information (threats) than information requiring mental calculations (fuel and TOT).

Lastly, we argue that the ideas and findings presented in this document have relevance beyond the simulated environment of the experiment. While this experiment focused on in-flight replanning for a military combat mission, the findings can be applied to many time-critical applications where automation integrates information to present a solution to the user. If we look at the experiment as an abstraction, we can see how the results may be generalized. The in-flight replanner display used in the experiment was simply the human-computer interface for accepting, rejecting, or modifying the automated solution. The in-flight replanning was a time-pressured decision-making task, with collaboration between a decision-support system and the human. In general, the experiment framed a time-critical task around the human-automation relationship.

[Page Intentionally Blank]

7. CONCLUSIONS

Chapter 7 highlights the important key findings and conclusions from our experiment. We also recommend future research for decision-support systems that would benefit the naturalistic decision-making field, and give a vision for the future use of the experimental results. The results from this experiment have the potential to influence current and future development of in-flight replanners, and more generally, pilot decision-support systems for conventional military aircraft, attack helicopters, and commercial aircraft. We were able to achieve the experimental research goals: to objectively and subjectively determine replanning performance as a function of automation assistance and time pressure, and develop a model for decision-support tasks.

- The experimental results demonstrated that we are able to quantitatively relate automation assistance and time pressure to replanning performance in a complex and time-critical environment using a route cost performance metric.

The results objectively showed that automated integration of information does assist in time-critical decision-making for the replanning task. Route automation assistance integrated one or a combination of hazard, or time-on-target (TOT) and fuel constraint information. Overall, subjects performed significantly better with route automation assistance that integrated both hazards and constraints (Full), and only hazards (Hazard), than without any automation assistance (None). Route cost trends suggested that replanning performance followed a non-monotonically increasing relationship with relaxing time pressures, with a significant increase in performance from 20 seconds to time scales greater than 20 seconds.

- With unlimited time, subjects outperformed the route automation assistance in every condition.

In our experiment, a subject always had information available to him or her that the route automation assistance did not have when suggesting a route. For example, while Full suggested a very good route, it did not process all the threat information available to the subject. Overall, unassisted subjects performed significantly better without time pressures than in all time-pressured conditions with automation assistance, even when a subject had Full and 55 seconds.

As hypothesized, subjects performed better than the route initially suggested by route automation assistance given enough time.

- Performance trends suggested a decreasing temporal benefit from route automation assistance.

The quantitative results suggested that the temporal benefit from each type of automation assistance decreased as available time increased, with greatest benefit at higher time pressures. The time at which automation assistance was no longer a benefit over None, or "characteristic time," depended on the type and degree of information integration. Using a linear model, the characteristic times were approximately 20, 55 and 160 seconds for Constraint, Hazard, and Full, respectively. The linear metric suggested that at extreme time pressures, any automation assistance was better than None. In addition, benefit from Full was greater than the sum of its Hazard plus Constraint module benefits at each time pressure, which supported the idea of compounding benefits from automation that more fully integrates information in suggesting a solution.

- The use of decision-aids that partially integrate information is not always beneficial.

Overall replanning performance with route automation assistance that only integrated constraint information (Constraint) was at best not significantly different than with None. The findings suggested that Constraint was actually a *detriment* to subject performance at time scales greater than 20 seconds. Similarly, route cost trends suggested that at 55 seconds, performance with None was at least as good as with Hazard. In addition, mission failures with Constraint (8 failures) and Hazard (13 failures) were higher than with None (6 failures), with Hazard having the most mission failures. These results support the idea that the design of decision-aiding automation must take a human-centered approach; simply automating because technology allows it may not always be appropriate.

- Performance trends suggested a non-monotonic increase in replanning performance.

Trends within objective results strongly suggested that at certain time pressures, subjects felt they had enough time to stray far from the automated initial route suggestion to improve route cost; consequently, however, subjects did not have adequate time to complete their task. From 28 to 40 seconds, average route cost increased. Mission failures more than doubled from 40 seconds (4 failures) to 55 seconds (11 failures). Finally, we observed a significant increase in the route modification rate with Full from 28 and 40 seconds (0.11 per second average) to 55

seconds (0.18 per second); however, route cost with Full did not significantly improve with time. Certain time pressures compelled subjects to remain cautious, and replan the suggested route locally; while lower time pressures appeared to adversely motivate subjects to replan more globally, with lower performance given more time.

- Decision-support systems should use adaptive automation for the integration of information from complex, uncertain, and time-critical environments.

Adaptive automation for decision-aids seeks to present the user with the most effective solution using the best types and amount of information, in direct response to the individual's needs and the demands of the dynamic and uncertain external environment. The decision-support system would determine the time pressures from current knowledge about the environment and situation at hand. For example, if a missile launched at an aircraft, the automation would need to calculate the appropriate decision-making time horizon. Applying the results from our experiment, an adaptive automation design could use the automation benefit metric. For example, in some situations, tasks warranting only the Hazard module may cease to suggest a route when it determined the time pressure to be more relaxed than at 55 seconds.

- This formal, yet abstract experiment is important to the design of decision-aids.

Many time-critical applications could benefit from the ability to quantify human performance through automation and time-pressure interactions. Literature in the naturalistic decision-making field echo this sentiment that empirical results from naturalistic environments are needed to better design decision-support systems [Cannon-Bowers, et al., 1996]. In life or death decisions, the appropriate type and amount of automated information integration are imperative. Adaptive automation may best suit in-flight replanning, especially in the complex and time-critical combat environments. While difficult, it is possible to design for adaptive and intelligent automation using objective metrics; we have taken the first step in this process.

7.1 The Future

We have several suggestions for future research and development. The results from this experimental study could be extended to include longer time scales. The motivation for this would be to determine with significance the characteristic times for each automation assistance category, particularly with Full automation assistance. Another experiment could develop a

flight simulation to quantify replanning performance specifically for military combat scenarios, using only military pilots. The simulation could also extend to real-time flight dynamics, instead of the static in-flight replanning environment we used. This flight environment would better simulate the dynamic and uncertain flight environment where updates may occur in the middle of replanning the current situation. The results from this experiment could more directly apply to developing technology.

Figure 7-1 illustrates our proposed decision-support model, which uses the same collaborative decision-making model as described in Section 3.1, In-Flight Replanning Overview. When deemed necessary, the decision-support automation (i.e., in-flight replanner) will use a prior performance information, specific to each user, to assist the user most effectively in time-critical situations. The a priori performance grid shown in Figure 7-1 breaks down how the system decides what would be the most effective automated support. The system first classifies the given scenario and rates the scenario difficulty, a design issue not addressed in our research. Each box within the scenario description grid has the a priori information on human decision-making performance, which varies by degree and type of automation assistance and time pressure. The performance results would come from human experiments designed to determine performance varying automation levels and time pressures at various task conditions and condition difficulties. For example, the experimental results described in this document may be used for one box of the performance grid, as shown in Figure 7-1.

Future in-flight replanning technology development could use this system model as a framework. The adaptive automation would use quantitative performance results, specific to an individual user, as the metric for what and how much information to integrate when suggesting solutions. Let us describe a vision for the future of in-flight replanning technology:

> A military pilot sits at a computer terminal, and runs through a brief experiment. The experiment evaluates quantitatively the pilot's performance in various scenarios, at conditions combining different automation levels and time pressures. Then on the flight line, the pilot engages the decision-support system, which includes an in-flight replanner module. The technology running behind the in-flight replanner displays has been configured previously to the performance abilities of that specific pilot, as determined from the experiment. Now, should a

time-critical situation arise in-flight, this adaptive technology will greatly enhance the decision-making capability of the human pilot by most optimally selecting how much and what type of automation assistance to provide. This technology will improve the likelihood of mission success in the face of an uncertain future.

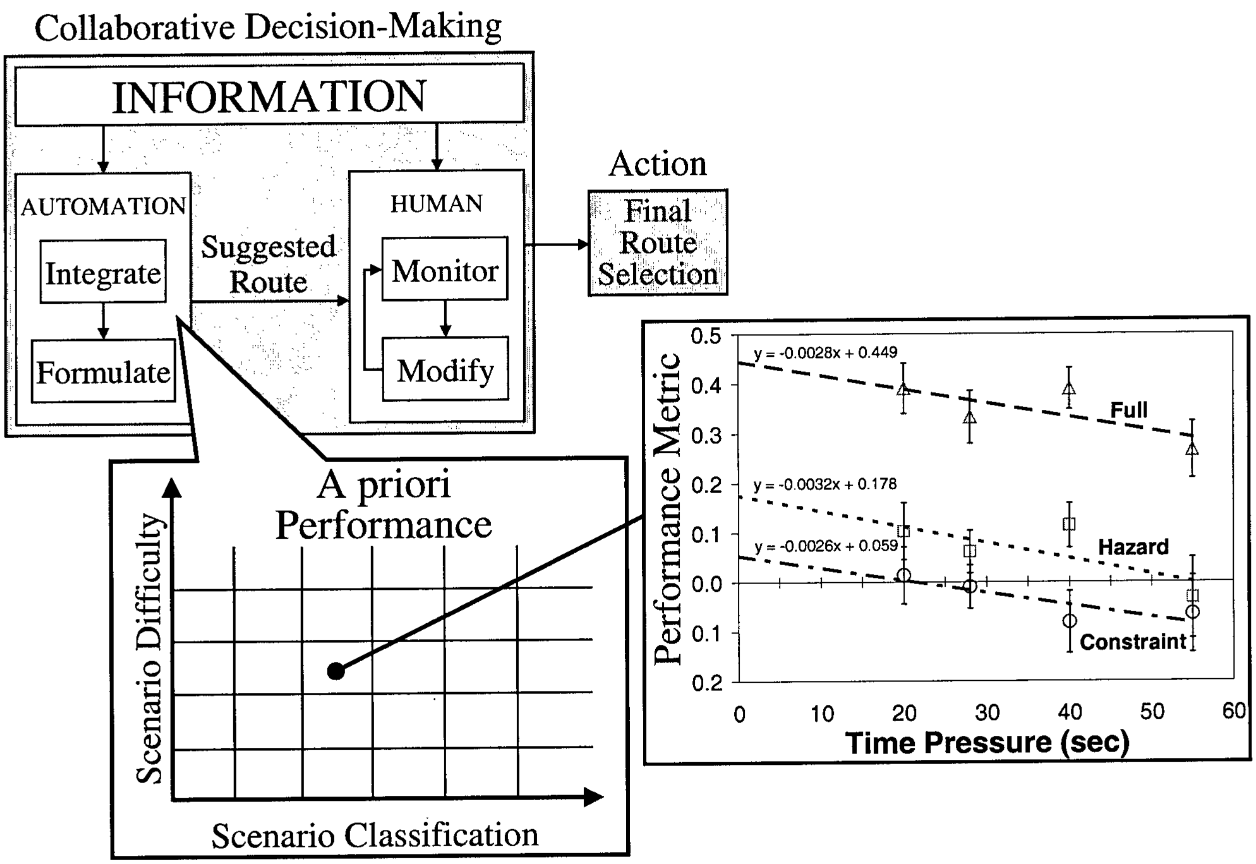


Figure 7-1. Proposed Model for In-Flight Replanner Technology.

User-centered automation design takes advantage of a human's intuition, experiences, and unique ability to process information not constrained by formal logic or rules. In the ever-increasing complex and lethal nature of combat environments, the need is evident for in-flight replanning cockpit technology, which can accurately and quickly assist pilots in making time-critical and life-dependent decisions. A pilot-centered approach to the design of decision-support automation is crucial for the successful implementation of this in-flight replanning technology.

[Page Intentionally Blank]

APPENDIX A (Consent Statement & Questionnaire)

INFORMED CONSENT STATEMENT

Experimental Study of Information Integration in Decision Aids for Planning Tasks

Participation in this study is voluntary and you may halt the experiment at any time and withdraw from the study for any reason, without prejudice.

This study is designed to evaluate the potential benefits of automated decision support in an in-flight replanning task. You will be monitoring a simulated aircraft navigation display (on a computer workstation) and making route modifications using the computer mouse and keyboard. You will be scored on your ability to accurately and rapidly develop flight plans that meet criteria for fuel burn, time, mission success, and hazard avoidance. There will be a brief questionnaire at the end of each test run. All data will be collected in a confidential manner and will not be linked in any way to your identity. You will remain anonymous in any report that describes this work. The study will take no more than 4 hours to complete, including breaks.

As with any use of computers, you may or may not experience headache, eye, neck, back, arm or hand strain, or fatigue. The experimenter will be stationed next to the computer, and frequent rest periods will be provided. Please inform the experimenter at the first sign of any uncomfortable symptoms, and the experiment will be interrupted and an attempt made to alleviate the cause of the symptoms. Should you wish to stop or delay the experiment, you are free to do so at any time. Also, please feel free to ask any questions or request clarification at any point in the study.

In the unlikely event of physical injury resulting from participation in this research, I understand that medical treatment will be available from the M.I.T. Medical Department, including first aid, emergency treatment and follow-up care as needed, and that my insurance carrier may be billed for the cost of such treatment. However, no compensation can be provided for medical care apart from the foregoing. I further understand that making such medical treatment available, or providing it, does not imply that such injury is the Investigator's fault. I also understand that by my participation in this study I am not waiving any of my legal rights.*

I understand that I may also contact the Chairman of the Committee on the Use of Humans as Experimental Subjects, M.I.T. 253-6787, if I feel I have been treated unfairly as a subject.

I volunteer to participate in this experiment, which is to involve making simulated in-flight replanning decisions on a computer workstation. I understand that I may discontinue my participation at any time, and that all data will be collected in a confidential manner and I will remain anonymous in any report that describes this work. I have been informed as to the nature of this experiment and the risks involved, and agree to participate in the experiment.

____________________________ ____________________________

Date Signature

* Further information may be obtained by calling the Institute's Insurance and Legal Affairs Office at 253-2822.

Questionnaire

Experimental Study of Information Integration in Decision Aids for Planning Tasks

Subject #____: Date:________

All data will be collected in a confidential manner and will not be linked in any way to your identity. You will remain anonymous in any report that describes this work. Your participation is voluntary and you may decline to answer any question on this questionnaire, without prejudice.

Note that this questionnaire contains both a written and an online component.

Subject Background

Age:_________Sex:_________

General computer experience (none, general, extensive):_____________

Piloting Experience (flight hours, ratings):_________________

Experience with flight management systems, if any:________________

To what degree were you able to replan within the time pressure?

1. Completed an optimal route
2. Completed a near optimal route, minimal improvement needed
3. Completed a good route, some improvement needed
4. Significant route improvement needed
5. Unable to improve or complete route

Scenario	Best				Worst
1	1	2	3	4	5
2	1	2	3	4	5
3	1	2	3	4	5
4	1	2	3	4	5
5	1	2	3	4	5
6	1	2	3	4	5
7	1	2	3	4	5
8	1	2	3	4	5
9	1	2	3	4	5
10	1	2	3	4	5
11	1	2	3	4	5
12	1	2	3	4	5
13	1	2	3	4	5
14	1	2	3	4	5
15	1	2	3	4	5
16	1	2	3	4	5

To what degree did the suggested route assist you in the replanning task?

1. Relied completely on suggested route
2. Suggested route provided great assistance
3. Suggested route provided some assistance
4. Suggested route provided very little assistance
5. Suggested route provided no assistance

Scenario	Completely			Not at all
1	1	2	3	4	5
2	1	2	3	4	5
3	1	2	3	4	5
4	1	2	3	4	5
5	1	2	3	4	5
6	1	2	3	4	5
7	1	2	3	4	5
8	1	2	3	4	5
9	1	2	3	4	5
10	1	2	3	4	5
11	1	2	3	4	5
12	1	2	3	4	5
13	1	2	3	4	5
14	1	2	3	4	5
15	1	2	3	4	5
16	1	2	3	4	5

To what degree did each of the following elements restrict your ability to develop an optimal flight plan?

Scenario	Element	Not at All	Slightly	Somewhat	Significantly	Completely
1	Time Pressure	1	2	3	4	5
	Hazard	1	2	3	4	5
	ToT	1	2	3	4	5
	Fuel	1	2	3	4	5
2	Time Pressure	1	2	3	4	5
	Hazard	1	2	3	4	5
	ToT	1	2	3	4	5
	Fuel	1	2	3	4	5
3	Time Pressure	1	2	3	4	5
	Hazard	1	2	3	4	5
	ToT	1	2	3	4	5
	Fuel	1	2	3	4	5
4	Time Pressure	1	2	3	4	5
	Hazard	1	2	3	4	5
	ToT	1	2	3	4	5
	Fuel	1	2	3	4	5

... for all 16 scenarios.

- To what degree did the training tutorial prepare you for the experiment?

1	2	3	4	5
Not at all	Slightly	Moderately	Mostly	Completely

Explain:

- In general, what approach did you take in planning your route:

- ***General Comments***

Please provide any other comments about the task or automation tools that you used:

APPENDIX B (Training Tutorial)

Dynamic In-Flight Replanner Tutorial

Revision 1.4
Updated February 27, 2002
October 19, 2001
ICAT, MIT

Background to the Experiment

The program that you are about to use is a simulation of a generic dynamic in-flight replanner (DIR). The DIR lets the pilot visually navigate a pre-determined route by displaying essential nav-aid information, as well as various hazards such as weather, traffic, or terrain.

The goal of this experiment is to find a quantifiable relationship between time pressures, information elements, automation levels, and resulting decision performance.

In general, you will be presented with a series of 16 different missions. In each mission, you will route-replan under time pressures to avoid in-flight hazards, meet time-on-target (ToT) goals, and satisfy fuel constraints. You will plan a 2-dimensional lateral route to arrive at a target point, via a rendezvous point, within a pre-defined time window, and to exit at the finish point with enough fuel. The mission is restricted to constant altitude and constant speed flight of a conventional aircraft.

Your performance will be evaluated by the quality of your route at the end of the time pressure, based on meeting route constraints and minimizing threat exposure and deviations from an optimal ToT.

This tutorial will explain the subject-computer interface and the experimental protocol. The training goal is to get you to steady-state and optimal performance for the data collection runs. **Follow the tutorial closely. DO NOT** exit the training to begin the data collection runs until I give you the OK. *Italicized* comments are action items for you to perform.

The Opening Screen

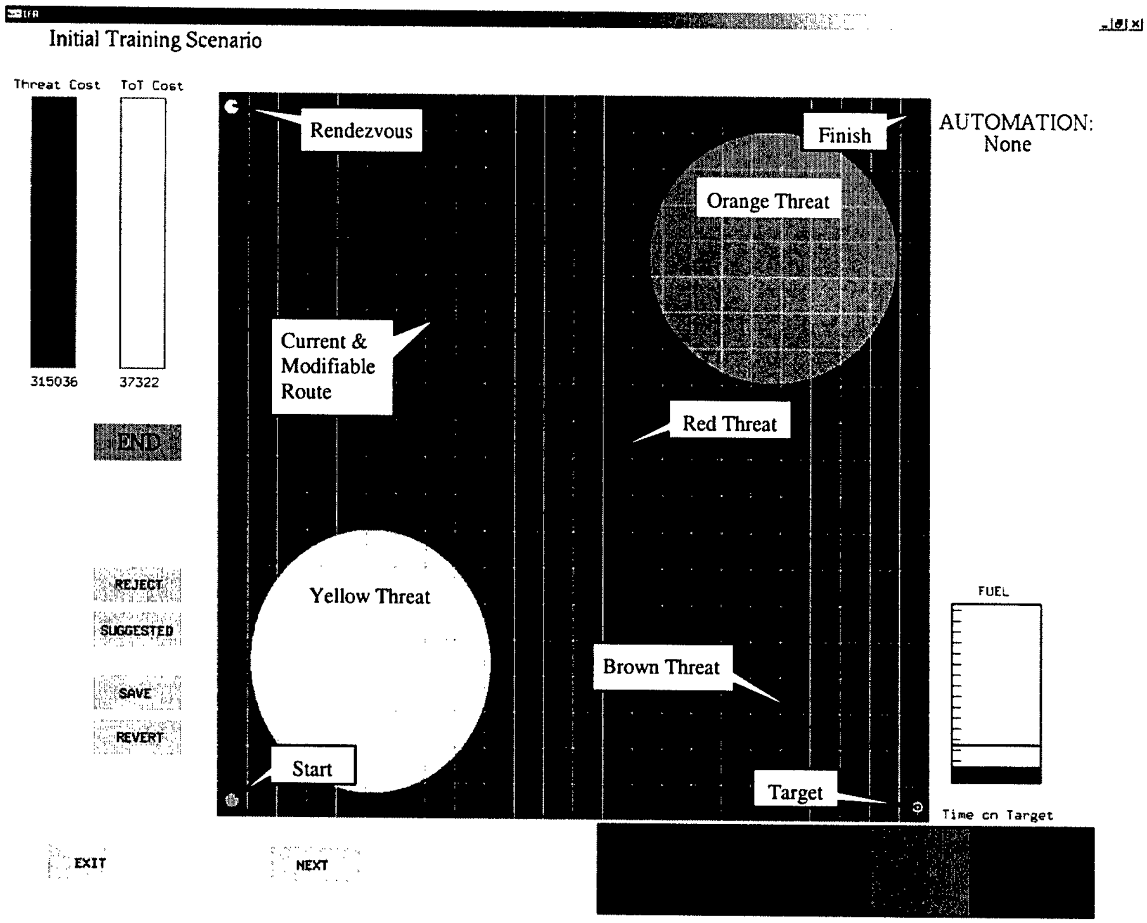
The opening screen is a dark green screen with a grid in the center of the window, and two buttons on the left. The grid indicates where the map will be shown. *Press the "space bar" to toggle the color scheme, choose what you like.*

The "SCENARIO" button will begin the series of 16 scenarios. This button may not be available if the simulation is configured as a trainer only.

The "TRAIN" button will bring up the training scenarios.

Press the "TRAIN" button.

The Initial Training Scenario



The Route

The bright blue line is the current and modifiable route. In correct order, the route begins at the **start point** (green dot), passes through the **rendezvous point** (white dot) and **target point** (blue dot with yellow crosshairs), and ends at the **finish point** (red dot). We will call these four points the "must-fly" points. In all scenarios, the route must hit each must-fly point in the above order.

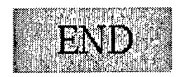
A route segment or an existing waypoint will highlight green when the mouse points to them. To modify the route, simply left click when the desired segment is highlighted anywhere along the route to create a new waypoint. Waypoints define the route's shape, and are blue diamonds. Once a waypoint is created, drag it to modify the route as desired. Right click on a waypoint to delete it. Although not displayed, a modifiable waypoint exists at each of the must-fly points. Note: Cannot delete the waypoint at the start point, and 2 waypoints is the minimum for a route.

Exercises

Add a waypoint anywhere along the route. Drag it around the screen, then delete it. Notice the highlighting of the route and waypoints.

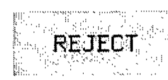
Modify the route to bypass the two brown circles in the upper-left and lower-right of the map, and then continue with the tutorial.

The Buttons



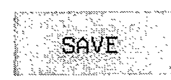
These buttons are located to the screen's left, and are route modification functions. They will highlight when pointed at, and will remain highlighted if selected.

The "END" button indicates final route acceptance. This button will end a scenario and take you back to the original screen.



The "REJECT" button clears all waypoints in your current route in excess of the 4 route-points, snapping to the route-points in the correct order.

The "SUGGESTED" button will return the current route to the computer-suggested route. If there is no suggested route, default is the original route.



The "SAVE" button will save the current route for future use.

The "REVERT" button will revert back to the 'save' route. If there is no saved route, default is the original route.

Exercises

Press the "SAVE" button. Now hit the "REJECT" button, what happens? Press the "REVERT" button, what happens?

Remove a waypoint from any must-fly point, and then press "END." What happens? Return the waypoint to the appropriate must-fly point. Notice how the waypoint snaps to the must-fly point.

Subject Performance

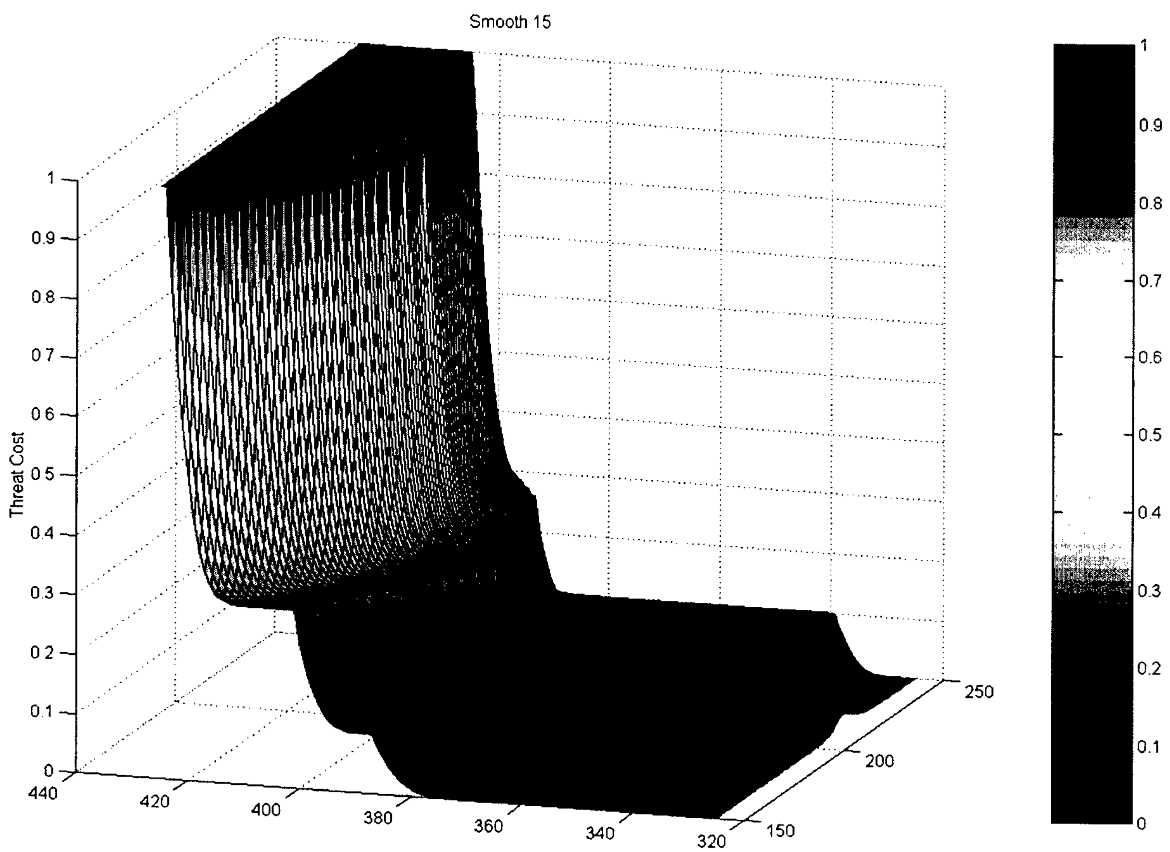
Subject performance will be measured by the route quality at the end of the time pressure. The route must be complete and acceptable, and minimize costs. The route cost function is composed of two parts, a threat and ToT deviation cost. The total cost is simply

$$\mathbf{Cost_{Total} = Cost_{Threat} + Cost_{ToT}.}$$

Threats

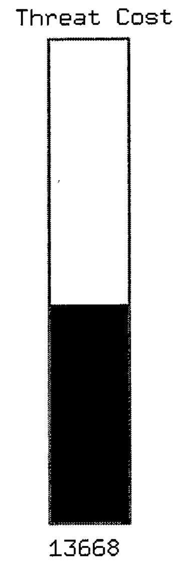
There are four threat levels, in increasing severity, indicated by yellow, orange, red, or brown colored shapes. If the route passes through a threat, then a linear cost is incurred. The longer the route intercepts a threat, the greater the cost.

Additionally, the DIR's automation can only detect and display approximate threat edges. To capture this element in the experiment, you will incur a cost the closer you push the threat edges. Below is a 3-d visual of normalized cost (vertical axis), as a function of (x,y) map coordinates. The cost figure is from a hazard cutout, with red, orange and yellow threat fields. You can observe how the cost is "smoothed" between the different threat cost levels. The smoothing will be a constant width throughout the experiment.



Cost rises smoothly and rapidly as the route approaches the next higher threat level. The cost ratio of brown, red, orange, and yellow threat is **50.0 : 1.0 : 0.3 : 0.1** . **Hitting a brown threat is NOT ACCEPTABLE!!!**

Threat Cost Gauge



The threat cost gauge is on the left side of the screen. It shows the route cost associated with passing through the map's hazard field. The bottom of the gauge digitally displays the cost, while the vertical blue bar provides a graphical representation. The maximum range is 30,000 for the vertical blue bar.

Note: Cost gauges will NOT be given in the data collection runs!

Exercises

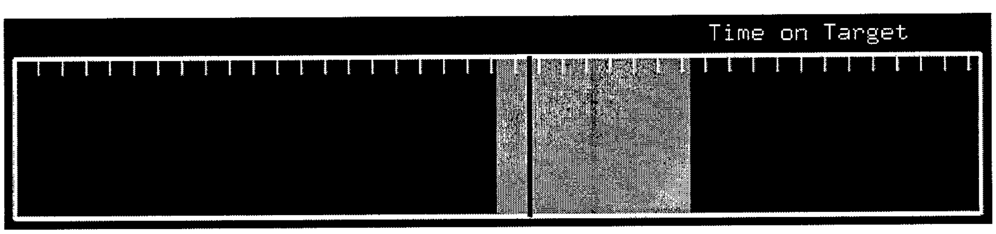
Modify the route so that it does not pass through any brown threats. Notice how much the cost has decreased. Now modify the route so that it barely skims the brown threat. Again, notice how sensitive the threat cost is to passing through a brown threat.

Now modify the route so that it passes through the red threat only. Modify the route until the threat cost due to the red threat is approximately 5,000. Repeat for the orange and yellow threats. Note how much longer the route must be in orange or yellow to incur the same cost as red. ***Commit to memory path lengths and their equivalent costs.***

Note: Use the map grid to estimate route length. Take advantage of this now, it will not appear in the data collection runs.

Modify the route to skim the edges of each color threat, one at a time. Note: This smoothing behavior will be the SAME for all scenarios. ***Commit to memory the smoothing behavior as you approach the edges.***

The Time-on-Target Gauge



Time-on-Target is the time to reach the **target** point from the start. In many missions, the pilot must reach the target within a certain time frame, or the mission is a failure. We have tried to simulate this aspect of mission planning by assigning a cost based on the difference between the actual ToT and the desired ToT.

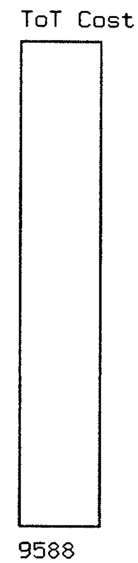
Reference the above picture. The ToT gauge shows a temporal relationship, with the left extreme representing time zero at the start point. The black vertical line represents the actual ToT associated with the current route. The green line represents the optimal desired ToT—what

you are aiming for. The blue area on either side of the green line represents the "acceptable range." The gauge will turn red if the actual ToT indicator moves outside the acceptable range.

In this particular case, the black line is to the left of the green line, which means that the current route reaches the target too early. To achieve a better ToT, you need to lengthen the route distance between the start and target points. You cannot adjust speed or altitude.

ToT Cost Gauge

The ToT cost gauge is on the left side of the screen. This gauge displays the digital cost at the bottom, while the vertical yellow bar provides a graphical representation. The maximum range is 30,000 for the vertical yellow bar.



The ToT cost reflects the difference between the desired ToT (the green line) and the actual ToT (the black or red line). Furthermore, there is an exponential cost within the acceptable range as the route deviates further from the desired ToT. **It is NOT acceptable to be outside the acceptable range, you will incur a severe cost penalty.**

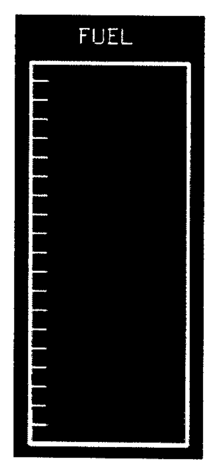
Exercises

Modify the segment of the route between the target point and the finish point. Do the ToT gauges change? Why or why not?

Modify the segment of the route before the target point. Watch how the actual ToT indicator moves real-time. How much do you have to increase or decrease the route length to make the indicator move one-tick mark on the gauge? ***Note, this behavior will be the SAME for all scenarios.***

What is the ToT cost at the boundaries of the acceptable range? What is the ToT cost when you hit the green line? What is the ToT cost in the middle of the acceptable range? Notice the exponential growth in cost within the acceptable range. ***Commit to memory ToT costs.***

The Fuel Gauge



The top of the gauge is Full, while the gray region is Empty. You MUST have enough fuel to complete the mission. The black horizontal line is the predicted fuel level to complete the mission based on total flight path length alone. When the predicted fuel level reaches the empty region, the gray will turn red.

Exercises

Modify the route so that your predicted fuel level drops into the empty region. Notice the change in color. While still in the empty region, press "END," what happens?

How much do you have to increase or decrease the route length to make the fuel indicator move one tick mark on the gauge? **Note, this behavior will be the SAME for all scenarios.**

Decision-Making

The challenge is to minimize both the route's threat exposure and its desired ToT deviation, while meeting the fuel constraint. It is unlikely that you will be able to plan the "perfect route," or be able to avoid high hazards completely within the given time pressure. Instead, you will be forced to balance competing goals in order to minimize the route cost and meet route constraints.

Exercises

Modify the route so that the ToT indicator is about halfway between the desired line and the edge of the acceptable range. What is the ToT cost? Now modify the route so that it passes only through the yellow threat. What distance do you have to travel through the yellow threat so that the threat cost equals the ToT cost you just measured?

Do the same for the orange, red, and brown (if possible).

From your results, how willing are you to deviate from desired ToT results in order to avoid a brown threat? What about the red, orange, and yellow threats? **Commit to memory this tradeoff!**

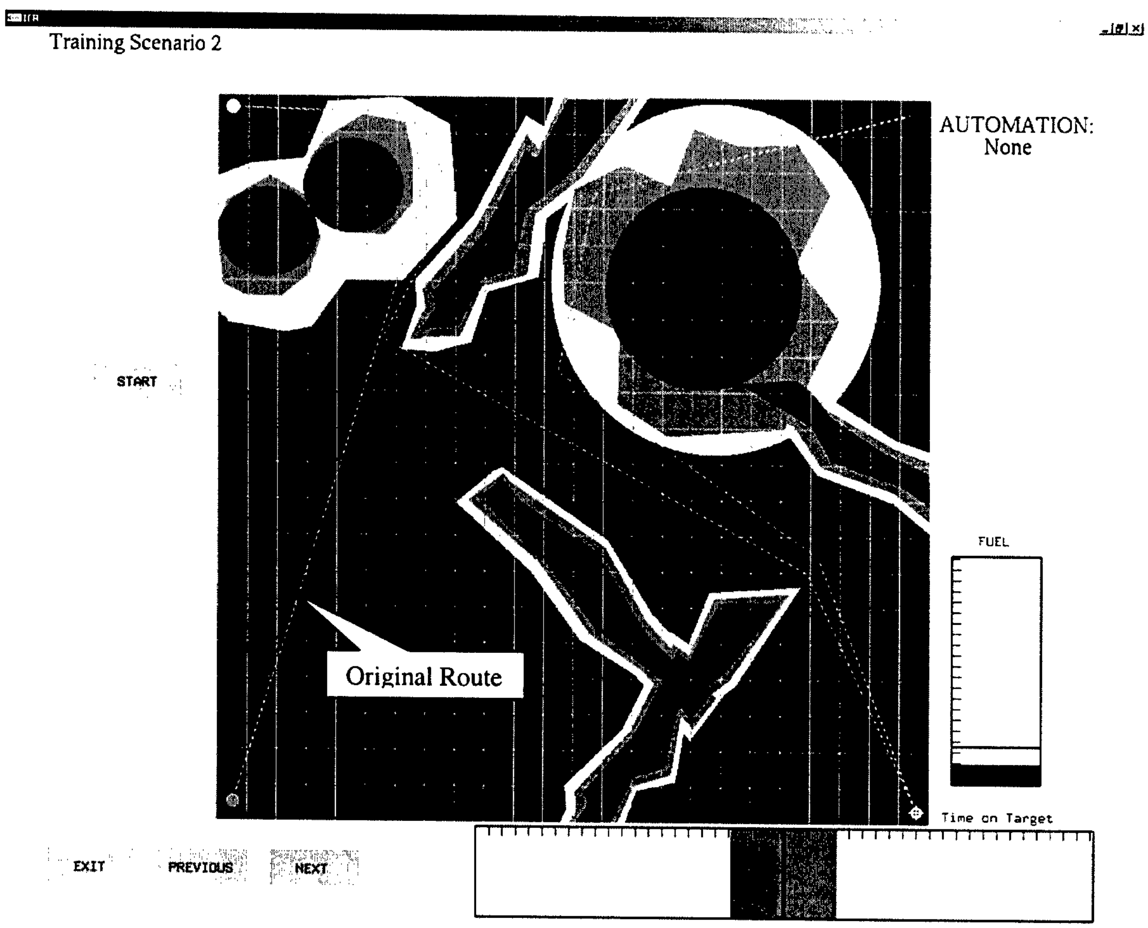
By now, you should understand the DIR interface, and also have a good feel for the threat and time costs. Next, you will learn the experimental protocol and further practice cost trade-offs.

Make sure that the route is valid (i.e. it passes through the must-fly points in the correct order, and the fuel indicator is above empty), and then press "END" to end the initial training scenario.

Experimental Protocol

READ FIRST!

After pressing the "SCENARIO" button, the screen will look similar to the following figure—this is the Pre-planned Mission. You can take as much time as you want to study the pre-planned mission. You cannot modify the original route. When ready (NOT NOW), you will press the "START" button to update the scenario.



The white dashed line is the initial plan. Since this is the preplanned mission, both the fuel and ToT constraints are satisfied, and the route avoids most of the threats.

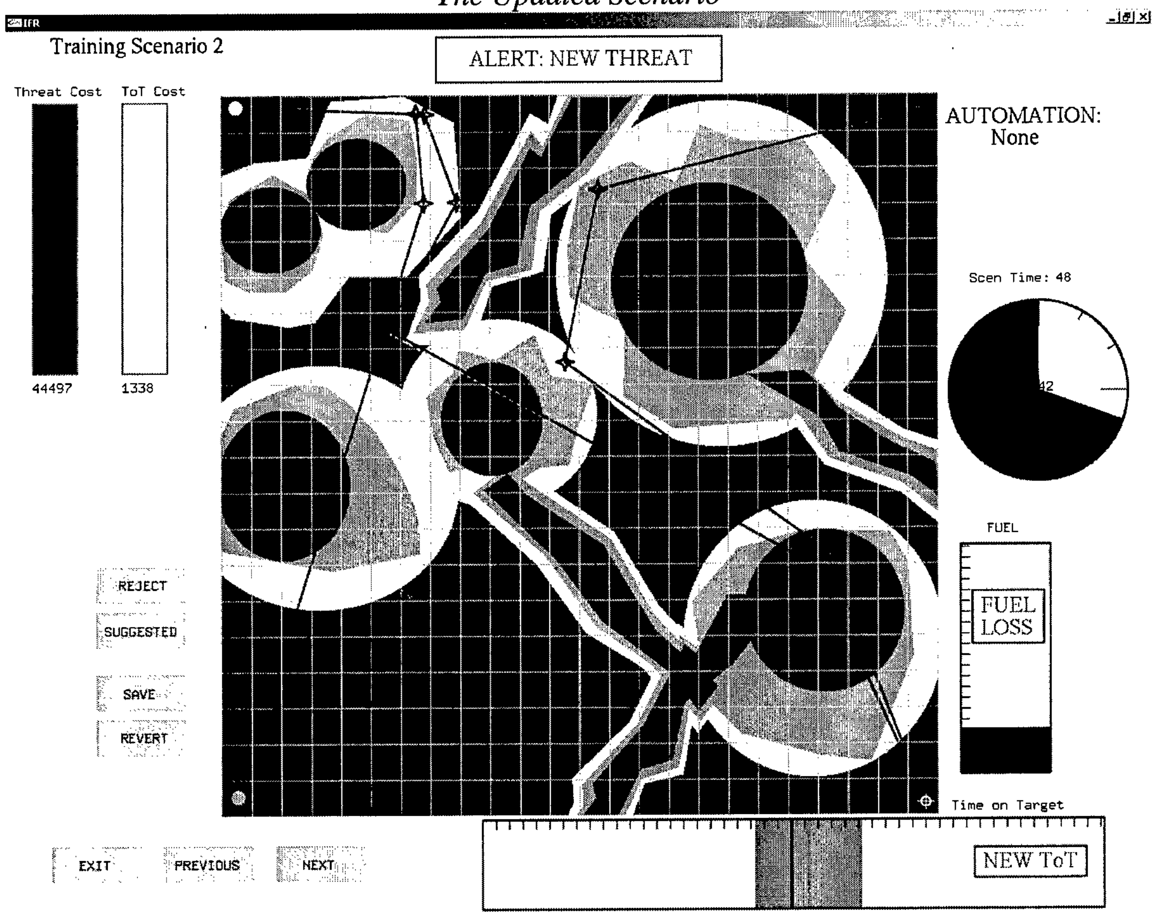
"AUTOMATION:" tells you the level of route automation you will receive for that mission. The automation will provide a suggested route that differs from the original route if there is a threat field, ToT, or fuel requirements change in the mission. There are four route automation levels:

- None—No additional suggested route will appear (the original route will be shown instead).
- ToT—The automation will relax the original route until it finds a solution that first optimizes your ToT, and then meets the fuel constraint. Threats are not considered. In the case that

you arrive too early, the automation will provide an appropriate length holding pattern at the Start point.

- Threat—The automation will relax the original route until it finds a solution that minimizes your threat exposure. However, the automation can only detect brown and red threats. ToT and fuel are not considered.
- Both—The automation integrates levels 2+3, first minimizing threat exposure, then relaxing the solution until ToT and fuel constraints are met. ToT will not be optimized with this automation, but ToT will be acceptable.

The Updated Scenario



A few seconds after pressing "START", the updated scenario will appear. Updates, with accompanying visual alerts, may include one or a combination of the following:

- A new threat (noted at the top of map).
- A new desired ToT (in ToT gauge).
- A loss of fuel, tighter fuel constraint (in fuel gauge).

The visual alerts will disappear when you first modify the route.

Again, the current and modifiable route is blue-colored. When there is no route automation, as in this example, the modifiable route initially mirrors the original route. When there is a computer-suggested route (automation 2, 3, or 4), the current and modifiable route initially mirrors the new suggested route. The computer-suggested route will remain as a dotted (or solid) magenta line; reference it as necessary. The original route will remain only in the scenarios with no automation.

A clock appears above the fuel gauge. The time available to replan is digitally displayed in the clock's center, while the red shaded area gives a graphical representation. The time elapsed since the scenario's start is digitally displayed above the clock. There will be an audible "beep" at the start, and at 5 seconds to the end of the countdown. In the data collection runs, you will face time pressures of 125, 90, 55, and 20 seconds.

****

Primary goal: you MUST have a <u>COMPLETE</u> and <u>ACCEPTABLE</u> route at the end of the time pressure.

****

To be complete and acceptable must satisfy each of the following constraints:

1. Meet fuel constraint
2. Meet ToT constraint
3. Avoid Brown threats
4. Hit all mustfly points

Secondary goal: minimize the route's cost to the best of your ability.

5. Minimize red, orange exposure & ToT deviation
6. Minimize yellow exposure

To reinforce the necessity of having a complete and acceptable route, the scenario will automatically give you a "feedback screen" at the end of the time pressure. The last completed route is used for all performance parameters. The feedback screen will show whether or not your final route was complete and acceptable. Accordingly, the following constraints will be marked either "Satisfied" or "Not Satisfied":

- Fuel constraint
- Mustfly points
- Brown threat
- ToT constraint

The route FAILS if "Not Satisfied" appears!!!

Press the "Continue" button to move to the scenario selection screen. For a few scenarios, however, pressing "Continue" brings you back to the same scenario at the point where it ended. For these scenarios, we are interested in your ability to optimize the route under no time pressures and given no route-automation support. Only when you feel you can no longer improve the route, press "END" to end the mission. Pressing "END" before you have completed an acceptable route will warn you of what is not satisfied. It is imperative that you give your best effort for this final optimization portion because it is an important parameter for data analysis.

Before you move to the next scenario, please fill out the questions in the questionnaire for the scenario you just finished.

Final Practice Runs

Finally, go through several example experimental runs. Note: Cost gauges will NOT be given in the data collection runs! *During the practice runs, cost gauges may be toggled on/off by pressing the 'q' key.*

EXIT PREVIOUS NEXT

These buttons, located to the screen's bottom left, allow you to navigate through the training scenarios. "EXIT" will take you out of training altogether. "PREVIOUS" and "NEXT" move through the training scenarios. *Use them as desired, but make sure you run each training scenario.*

A few tips...

1. Take time to preview mission before you press start button
   Identify start, rendezvous, target, and finish points
   Make sure you know the time pressure: 125, 90, 55, or 20 seconds.
   Make sure you know the automation
2. The time pressures will be high; solving route segment by route segment will not work in most cases! You must take a **GLOBAL** approach to making a complete and acceptable route, at least initially. Does this route intercept brown hazards anywhere? Is there enough fuel? Is my ToT acceptable? Fix the constraints first! You may miss these solving the route segment by segment. Then minimize the threats.
3. Sometimes, it may be easier to hit "REJECT" initially.
4. As the time pressure ends, make sure the route is complete and acceptable before you release the mouse button. Pushing the fuel/ToT boundaries near the scenario's end is dangerous!
5. Do not try for big changes when you have no time to verify!
6. Do not push the brown threat boundary!
7. ToT automation may change the finish segment of the route to meet the fuel constraint...check for any brown threat intercepts!
8. You may want to trade fuel on the last segment for both ToT and Threat costs before the target.

Scenarios 1-2

In these training scenarios, you will not be given any automation for route planning. The purpose of this training is to develop your own route cost minimization strategies, and to further refine your ability to balance competing interests. In Scenario 1, you can take as much time as you like to replan the route. Scenario 2, however, will introduce a time pressure. Learn to use the route edit buttons. Press "REPEAT" to repeat any of the training scenarios. Pressing "END" will bring you to the next training scenario.

Scenario 3

This scenario introduces route-planning automation that minimizes threat exposure. However, the automation can only detect the two highest threat levels, brown and red. This automation does not guarantee an acceptable suggested route; it will be the lowest threat cost route. There is no filtration of fuel and ToT information with this automation. You may have to relax this route to meet ToT and fuel constraints. Learn how to use this route planning automation to your advantage!

Scenario 4

This scenario introduces ToT route-planning automation. With the original route, the automation first optimizes for ToT and then satisfices the fuel constraint. The automation will implement a holding pattern at the Start point if the route arrives at the target too early. This automation does not guarantee a low cost route; it will be an acceptable route. There is no filtration of threat information with this automation. The automation may alter the original route to now intercept brown threats, pay attention! Learn how to use this route planning automation to your advantage! Learn how to use the holding pattern to quickly adjust the fuel/ToT constraints.

Scenario 5

This scenario introduces route-planning automation that integrates threat and constraint information to produce a complete and acceptable route. The original route will be relaxed to minimize threat exposure, and then relaxed further to meet the ToT and fuel constraints. This automation will NOT be the lowest cost route; it will be a complete and acceptable route. Learn how to use this route planning automation to your advantage!

Scenarios 6-12

These scenarios will introduce you to the greatest time pressures of the experiment. Learn what is possible to do within these time pressures. Learn how to initially solve the problem GLOBALLY! Learn when to use the route modification buttons!

You should reach steady-state and optimal performance. *Navigate through the series of training scenarios as desired.* ***DO NOT exit the training to begin the data collection runs until I give you the OK.***

Ask me QUESTIONS!

APPENDIX C (Software Architecture)

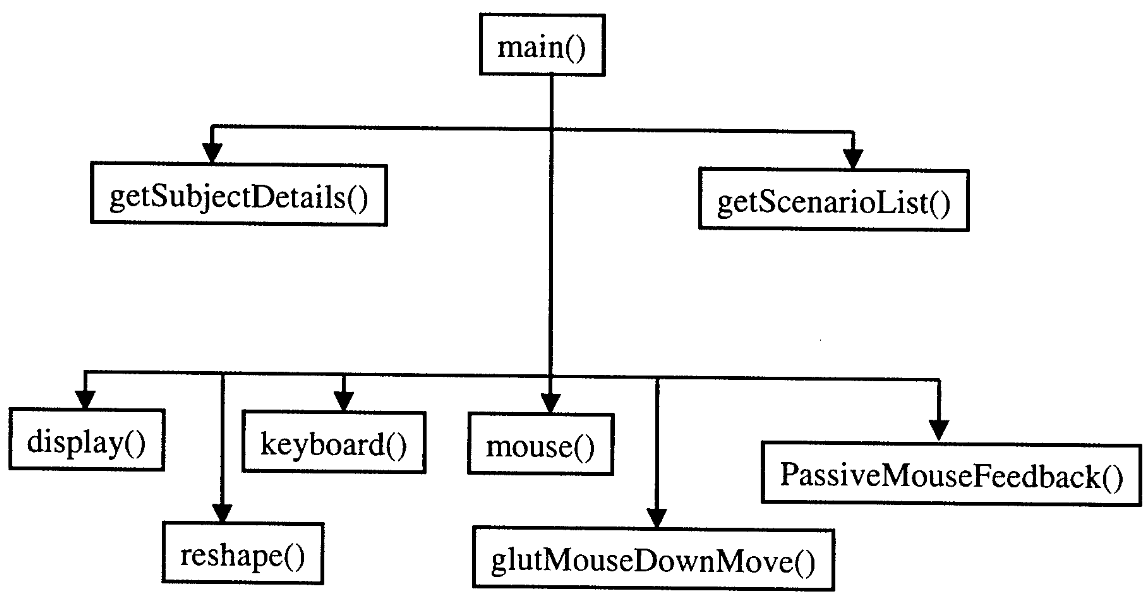
Farmey Joseph, my undergraduate research assistant during 2001, deserves full credit for the material in Appendix B.

Dynamic In-Flight Replanner (DIR)

The DIR (or In-Flight Replanner, IFR) was designed to carry out the following tasks:

- Input details about the experimental subject.
- Allow the subject to select scenarios to run, or (alternatively) input a list of scenarios from an external data file and present those scenarios to the subject.
- Provide an interface for the subject to carry out training and to move between scenarios.
- For each scenario, open and read 3 different data files that contain information about the scenario threats, automation, constraints, etc.
- From those data files, draw the scenario map, constraint gauges, alerts, etc. for both the old and updated scenarios.
- Allow the subject to alter the route on the updated scenario and observe the real-time effect on the fuel and time-on-target (TOT) gauges.
- Calculate the cost of the route, and re-calculate every time the user clicks a mouse button.
- Display a cost gauge (for training scenarios only).
- Record the route cost vs. time data, and print to an external data file at the end of each scenario.
- Record a summary of the costs for all the scenarios run by a subject, and print to an external file when the subject exits the program.

DIR Software Architecture Overview:



New Scenario -------------- initialize() -----------------→ readInit()
→
loadCostFile()
clearObjects()
ReadScenario()
InitializeCurrentPlan()

glutMouseDownMove()

- Search for highlighted waypoint (using currentplan.point[i].mode)
- Set highlighted point = mouse location
- Check for nearby mustfly points
- Snap if close

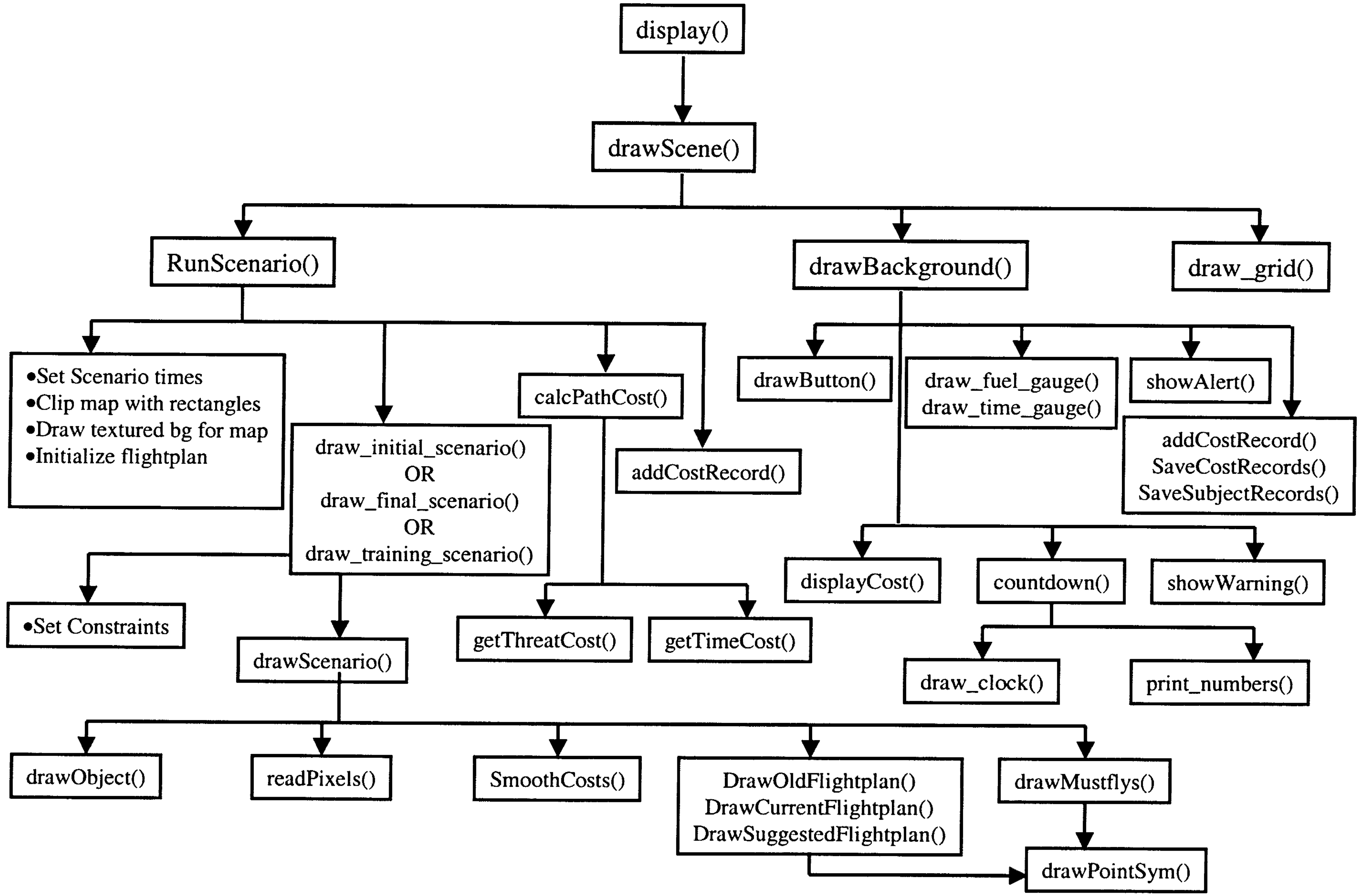
mouse()

- Button functions
- Add/Delete waypoints
- Update cost records on mouse-up

PassiveMouseFeedback()

- If mouse is over route, highlight that part of the route

DIR Simulation Details:

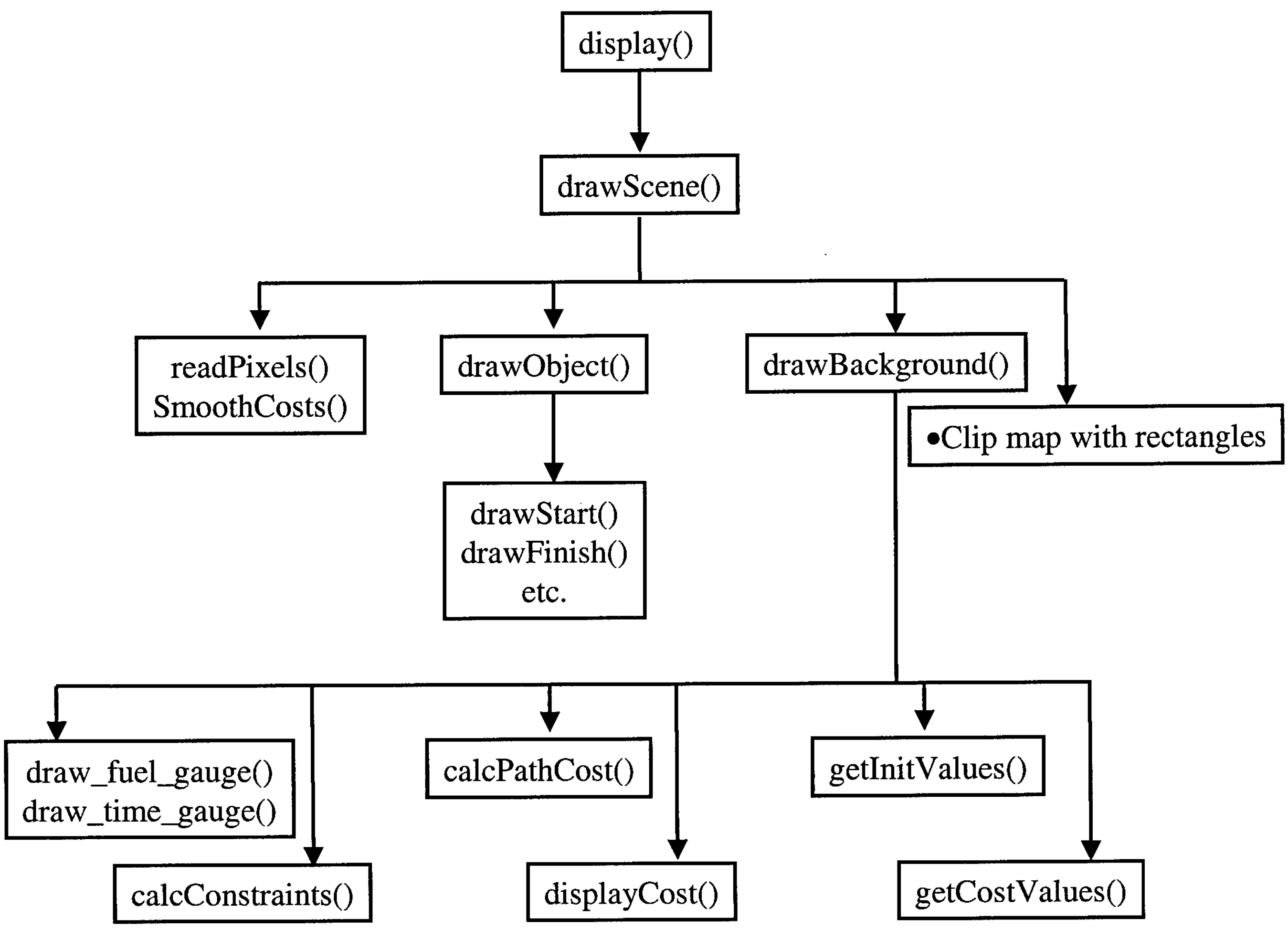


Overview of the Scenario Editor (SE)

The Scenario Editor was designed to carry out the following tasks:

- Provide all the features of a basic paint program (draw objects of different shapes & colors, move, copy, delete, undelete, load, save, clear, etc.)
- Distinguish between files for an old scene, a new scene, or a training scene.
- Allow the user to specify constraints for a scenario.
- Display constraint gauges.
- Calculate the cost of a route and display a breakdown of that cost.
- When displaying a "new" scene, incorporate information from the corresponding "old" scene in order to display the old route and to calculate the constraints.

Scenario Editor Software Architecture



Overlap between the DIR and the Editor

The Editor was originally designed as a paint program that could be used to draw the different scenes that would then be displayed using the IFR. We gradually realized that the process of creating scenarios was a lot more complex than we had thought. In order to create a good scenario in a reasonable amount of time, we found that it was necessary to have immediate feedback of cost and constraint information. To provide this feedback, we took portions of the IFR code and incorporated them into the Editor. We also made it possible to edit a route inside the Editor in the same way as in the IFR. The resulting overlap includes:

- Code that calculates the fuel and ToT constraints and draws the gauges: DistToMustfly(), draw_fuel_gauge(), draw_time_gauge().
- Code that calculates the cost of a route: ReadPixels(), SmoothCosts(), calcPathCost(), calcThreatCost(), calcTimeCost(), calcFuelCost().
- Code that allows the user to insert, delete, and move the waypoints of a route: mouse(), ActiveMouseFeedback(), PassiveMouseFeedback().

External Data Files

initX.txt

This file contains the countdown time, the fuel and ToT constraints, the type of alert, and the level of automation for Scenario X. It is created in the Editor program, and read by the IFR program at the beginning of every scenario. The init files for the training scenarios are designated as initXt.txt, where X is the number of the training scenario. The only exception is the initial training scenario, which uses the file init.txt.

cost.txt

This file contains the cost for a brown, a red, an orange, and a yellow pixel. It also contains the weightings for the threat cost and the time-on-target cost. Both the Editor and the IFR work off this file. The Editor loads the file every time the program is run, while the IFR loads it every time a new scenario is begun.

osceneX.txt
nsceneX.txt

These files contain the data for the objects (threats, mustfly points, and routes) for Scenario X. The oscene file contains the objects for the original scene in the scenario, and the nscene file contains the objects for the updated scene. For the training scenarios, these files are designated osceneXt.txt and nsceneXt.txt. For the initial training scenario, the objects are contained in the file tscene.txt. These files are created within the Editor and read by the IFR at the beginning of every scenario.

ground.bmp

This bitmap file contains the image that is used for the background of the HSI in the IFR. It is used only by the IFR.

Testing the Cost Function

To test the cost function, we added code to the function getThreatCost() that would print the coordinates of every point in the route, and the corresponding cost of each point, to an external data file named outdata.txt. (This code is normally commented-out.) We also un-commented the code that prints the coordinates of the mouse pointer whenever a mouse button is clicked. Then we created a scenario with four threats—one for each color. We ran it in the IFR and adjusted the route so that it passed through each threat. We then positioned the mouse of each intersection of the route and a threat and recorded the coordinates. Finally, we opened up the outdata.txt file and examined its contents. We checked that the route passed through the intersections we had recorded, and that the cost for the points between these intersections equaled the cost for the color of that threat. We also checked that the points outside the threats received a zero cost.

IFR Technical Details

5 source files: bitmap.c, cost.c, draw.c, filehandle.c, main.c
4 header files: bitmap.h, , cost.h, draw.h, main.h

<u>bitmap.c</u>

Functions:

```
GLubyte * LoadDIBitmap(const char *filename, BITMAPINFO **info)
```

This file contains the code for loading a bitmap file into memory. Most of the code was borrowed from Michael Sweet.

<u>cost.c</u>

Functions:

```
void smoothCosts(void)
void readPixels(void)
void addCostRecord(void)
void saveCostRecords(void)
void saveSubjectRecords(void)
void DisplayCost(int xpos, int ypos, int width, int height)
float calcPathCost(void)
float getTimeCost(vertices plan)
float getFuelCost(vertices plan)
float getThreatCost(vertices waypoints)
```

This file contains almost all the code used for calculating the cost of a route.

<u>draw.c</u>

Functions:

void DisplayAutomation (void)
void drawRect (int x0, int y0, int x1, int y1, fillType fillFlag)
void drawEllipse (int x0, int y0, int x1, int y1, fillType fillFlag)
void drawPolygon (vertices polygon, fillType fillFlag)
void drawCloud (vertices polygon, fillType fillFlag)
void drawObject(int i, int arraytype)
void drawScenario(int arraytype)
void DrawCurrentFlightplan(void)
void DrawOldFlightplan(void)
void DrawSuggestedFlightplan(void)
void DrawMustFlys(int scenetype)
void DrawPointSym(point2d)

This file contains the code for drawing most of the elements of the Horizontal Situation Indicator, including the routes and threats.

<u>filehandle.c</u>

Functions:

void initialize(void)
void readInit(void)
int readScenario(int scenetype)
void InitializeCurrentplan(int plan)
int loadCostFile(void)
void clearObjects(void)

This file contains the code for loading the scenario data from external files.

<u>main.c</u>

Functions:

void mouse(int button, int state, int x, int y)
void PassiveMouseFeedback(int x, int y)
void glutMouseDownMove(int, int)
void reshape(int w, int h);
void keyboard (unsigned char key, int x, int y)
void display(void)
void init(void)
void draw_button(int buttonname, float xpos, float ypos)
void draw_initial_scenario(float xpos, float ypos, float width, float height)

void draw_final_scenario(float xpos, float ypos, float width, float height)
void draw_training_scenario(float xpos, float ypos, float width, float height)
void draw_fuel_gauge(float xpos, float ypos, float width, float height, float ActualFuel)
void draw_time_gauge(float xpos, float ypos, float width, float height, float ActualToT)
void glprint(char text[], int text_length, float x_pos, float y_pos, int font_size)
void draw_clock(float time, float xpos, float ypos, float radius)
void print_numbers(int count, float xpos, float ypos)
void draw_grid(float xpos, float ypos, float width, float height)
void showWarning(void)
void showAlert(void)
void getSubjectDetails(void)
int getScenarioList(void)
void runScenario(void)
void countdown(clock_t scen_time)
void ClearFlightplan(void)
void ClearToStartFinish(void)
void ClearToMustflys(void)

This file contains the code for running the scenarios.

<u>bitmap.h</u>: header file for bitmap.c.
<u>cost.h</u>: header file for cost.c.
<u>draw.h</u>: header file for draw.c.
<u>main.h</u>: header file for main.c and filehandle.c.

Editor Technical Details

5 source files: cost.c, draw.c, filehandle.c, functions.c, IFR.c
1 header file: all.h

<u>cost.c</u>

Functions:
void DisplayCost(void);
float calcPathCost(void);
float getTimeCost(vertices plan);
float GetThreatCost(vertices waypoints);
float getFuelCost(vertices plan);
void smoothCosts(void);
void readPixels(void);

This file contains the code for reading in the pixel color data for a scene, assigning a cost to each pixel, smoothing the cost array, and then calculating the cost for a given route.

<u>draw.c</u>

Functions:

void mouse(int button, int state, int x, int y)
void reshape(int w, int h)
void display (GLvoid)
void ActiveMouseFeedback(int x, int y)
void PassiveMouseFeedback(int x, int y)
void keyboard (unsigned char key, int x, int y)
void arrow_keys (int a_keys, int x, int y)
void init()
void CleanUp(void)
void drawStart(int x, int y)
void drawFinish(int x, int y)
void drawTarget(int x, int y)
void drawMustfly(int x, int y)
void drawSteer(int x, int y)
void drawRect (int x0, int y0, int x1, int y1, fillType fillFlag)
void drawEllipse (int x0, int y0, int x1, int y1, fillType fillFlag)
void drawPolygon (vertices polygon, fillType fillFlag)
void drawCloud (vertices polygon, fillType fillFlag)
void drawRoute (vertices polygon, fillType fillFlag)
void drawObject(int i)
void drawScene()
void DrawOldFlightplan(void)
void makeMenu()
void colorMenu(int value)
void handleMenu(int value)
void routeMenu(int value)
void objectMenu(int value)
void objectFillMenu(int value)
void objectEditMenu(int value)
void fileMenu(int value)

This file contains the code for: starting and ending the program, initializing and updating the window, the editor menu system, drawing objects

<u>filehandle.c</u>

Functions:

int saveToFile(void)
int saveScenario(void)
int loadNewFile(void)
int loadFile(void)
int saveInitFile(void)
int saveCostFile(void)

This file contains the code for saving scenario and data files, and loading scenario files.

functions.c

Functions:

```
void deleteObject (int num)
int  selectObj (int x, int y)
int  clearDisplay (void)
void pushObject (Object list [], int i)
void popObject (Object list [], int i)
void flipHorizontal (int selectedObject)
void flipVertical (int selectedObject)
void undoDelete (void)
void copy (int selectedObject)
void rotate (int selectedObject)
void finishObject(void)
int selectPoint(int x, int y)
void sortObjects(void)
void flipThreatOrTerrain(int ObjSource, int FlipType)
```

This file contains the code for manipulating objects and for clearing the display.

IFR.c

Functions:

```
void drawBackground(void)
float DistToMustfly(vertices plan, int type)
void draw_fuel_gauge(float xpos, float ypos, float width, float height, float ActualFuel)
void draw_time_gauge(float xpos, float ypos, float width, float height, float ActualToT)
int loadInitFile(void)
void glprint(char text[], int text_length, float x_pos, float y_pos, int font_size)
void getInitValues(void)
void calcConstraints(void)
int getOsceneDistances(void)
int loadCostFile(void)
void getCostValues(void)
```

This file contains the code necessary for implementing the IFR component of the Editor—i.e., the code that calculates constraints and draws the constraints gauges, measures the length of a route, opens the corresponding oscene file (when loading an nscene file), and determines the cost parameters. It also contains a function that uses GLUT commands to print text to the screen.

APPENDIX D (Subject Comments)

Subject Comments from Post-Experimental Questionnaire/Interview

To what degree did the training tutorial prepare you for the experiment?

7.
wanted to see scenario again after given route feedback
8.
Felt there could be more emphasis on the trade off between excess fuel on last route leg for hazard costs
12.
wanted more training scenarios
13.
confused with route direction initially
14.
would have liked more practice
16.
wanted more scenarios, and to have my verbal automation tips written
17.

- wanted more scenarios with varying time pressures
- would have like option to make scenarios more difficult

Scores:
5 5 4 5 4 4 4 4 4 4 4 4 4 3, average = 4.1

The training tutorial "mostly" prepared subjects for the experiment. The general consensus for improvement would have been to include more training scenarios. It becomes much easier to critique the training when all is done. Thus, while most subjects wanted more training, given that there was a finite amount of time, they would agree the training adequately prepared them for the problem difficulty encountered in the data collection scenarios. Their comments are especially balanced since most subjects were very anxious to begin the data collection scenarios, being tired of training.

In general, what approach did you take in planning your route?

Initial observations are analyzing information contained in the pre-planned mission: the hazards, the route direction, the given automation assistance and what it would do to the route, predicting possible new hazards.

4.

- initial observations
- time awareness, then meet constraints, then optimize
- Given 20 seconds, subject just tried not to violate any constraints.

6.
- Under high time pressures, goal was to avoid brown hazards and have enough fuel.
- ToT information element was rarely a big deal.
- Given 55 seconds, subject felt their route was 80-90% optimized.
- Didn't bother with orange or yellow hazards.

7.
- initial observations
- Subject hit "revert" once to make route satisfactory.
- Subject gave the fuel constraint and brown hazards the most priority.

8.
- initial observations
- Given no time, would first check fuel constraint, and then avoid brown hazards.
- Given time, would meet fuel constraint, avoid brown hazards, and finally optimize red, ToT, orange.

9.
- With none or ToT/fuel auto assist, would watch for brown hazards.
- With hazard auto assist, would watch for ToT constraint.

10.
- The approach to replanning was very dependent on automation and time.
- Would only try for major reroutes if time permitted.
- Replanning depended on the scenario, was there enough fuel and was the route's ToT too early?

11 and 12. no comments

13.
- Initial observations
- Most of the time changes were made from the suggested route.

14.
- Felt that automation helped to predict problems.
- Would first quickly fix brown hazards.
- If the fuel constraint was a problem, subject looked for route segments that can be shortened
- If time permits, can I spend less time in red?
- Did not much worry of ToT information element.

15.
- Evaluated constraint violations and corrected them accordingly.
- Sometimes, the last minute changes were very costly.

16.
- Subject mostly relied on the suggested route.
- Pressed "revert" once.

17.
- Having more time got me focused on cost optimization, and forgetting about goals and that optimization may change constraints.

18.
- initial observations
- priority: fuel, brown, within ToT window

- Watched for holding pattern with ToT auto assistance.

Please provide any other general comments about the task or automation tools that you used:

4.
   - Subject felt hazard auto assist handy, only having to worry of fuel and ToT constraints.
   - Did not like ToT/fuel auto assist because it did not consider hazards and had to go back and forth between the holding pattern to get fuel if needed.
   - The ToT and fuel coupling was confusing, especially for the final route leg after the target.

6.
   - Subject simply tweaked the full auto assist suggested route.
   - Preferred ToT/fuel to hazard auto assist because it was easier to adjust for hazards than to adjust for ToT/fuel constraints.

7.
   - Felt that automation assistance reduced uncertainty, but not necessarily the workload.
   - Hazard auto assist was the least useful, requiring significant rerouting to meet fuel constraints.
   - ToT/fuel auto assist was a little better, requiring just brown hazard avoidance.
   - Full auto assist was the most helpful.
   - With no auto assist, the subject had no idea of the impending problem. With automation assistance, at least some of the problem was solved.

8.
   - Subject was better able to avoid hazards then meet the route constraints. Thus, ToT/fuel was preferred to hazard auto assist.
   - Felt there was a big advantage from ToT/fuel auto assist in that it shows you how much path can be used for avoiding obstacles.
   - Felt full auto assist the best, and approached these scenarios with "can I improve a little?" Subject felt like he or she made the route worse at times when given full auto assist.
   - No auto assist was the least liked.

9.
   - Subject liked having full auto assist.
   - Preferred to have no auto assist than both ToT/fuel and hazard auto assist because subject was free to look at the scenario and possible solutions. However, felt both ToT/fuel and hazard auto assist were somewhat useful.

10.
   - Liked having to only optimize with full auto assist.
   - Preferred ToT/fuel to hazard auto assist because subject felt adjusting to meet fuel constraints was difficult under time pressures, while just optimizing the visual hazards was easier.
   - Preferred none to hazard auto assist.

11. no comments

12.
   - Subject claimed that the overall goals were not followed all the time in trying to optimize cost.

- The suggested route was used as a starting point.
- Liked making small corrections to full auto assist.
- Subject felt partial auto assist helped in about half of their scenarios, at least being helpful in its title area.
- Preferred hazard to ToT/fuel auto assist because subject could work the route to meet constraints easier.
- None was the least liked.

13.
- Subject like full auto assist the best.
- Preferred ToT/fuel to hazard auto assist because subject knew how to fix the possible hazard problem.
- Preferred none to hazard auto assist because with hazard auto assist subject could not predict what the route would do.
- Subject liked having some scenarios that were nearly impossible to handle.

14.
- Felt it was impossible to trust ToT/fuel and hazard auto assist, and not sure if there was any benefit to having those automation categories.
- Preferred to have either none or full auto assist, than the partial automation.
- ToT/fuel auto assist slightly more preferred than hazard auto assist.

15.
- Subject relied heavily on full auto assist.
- The benefit to having partial automation depended on the scenario.
- Overall, subject tended to rely on suggested route.

16.
- Subject felt there was too much clutter, mentioning getting confused with the original/suggested routes.
- Would rather have some auto assist than none.
- Preferred ToT/fuel to hazard auto assist because subject felt it provided some useful tools.
- Like full auto assist.

17.
- At times, subject forgot about fuel constraints when there was time.
- Tended to start with the suggested route because subject felt there were always good aspects to the suggested route; subject never used "reject."
- Had a good feel for the problems that would occur having full and ToT/fuel auto assist- like these auto assists very much.
- Preferred having hazard auto assist than none.
- Was most fearful of the impending problems with no auto assist.

18.
- Tended to start with the suggested route.
- Subject felt that automation gives a perspective on the problem.
- Felt full auto assist was the best.
- Both ToT/fuel and hazard auto assist were slightly better than none.
- Subject mentioned that brown hazard, not red, as the highest level went against intuition

[Page Intentionally Blank]

REFERENCES

Aust, R (1996, May 20-23). "Mission Planner Evaluation in the Mission Reconfigurable Cockpit." IEEE National Aerospace and Electronics Conference, Dayton, OH. Vol. 1, A96-37576 10-1. Institute of Electrical and Electronics Engineers: New York. pp. 102-108.

Boeing Company (2002). "Air Traffic Management: Revolutionary Concepts that Enable Air Traffic Growth While Cutting Delays." WEB: http://www.boeing.com/atm

Cannon-Bowers, J.A., E. Salas, and J.S. Pruitt (1996). "Establishing the Boundaries of a Paradigm for Decision-Making Research." *Human Factors*, vol. 38 (2), pp. 193-205.

Cobb, G.W. (1998). Introduction to Design and Analysis of Experiments. Springer-Verlag: New York.

Colucci, F. (1999). "Flying with Experts: The Rotorcraft Pilot's Associate Shows How Expert System Can Manage Cockpit Information." *Vertiflite*, vol. 45 (4), pp. 18-21.

Dornheim, M.A. (1999, October 18). "Rotorcraft Pilot's Associate: Apache Tests Power of New Cockpit Tool." *Aviation Week & Space Technology*, pp. 46-49.

Endsley, M.R. and E.O. Kiris (1995). "The Out-of-the-Loop Performance Problem and Level of Control in Automation." *Human Factors*, vol. 37 (2), pp. 381-394.

Fan, T.P., D.S. Hyams, and J.K. Kuchar (1998, August 10-12). "Study of In-Flight Replanning Decision Aids." AIAA GNC Conference Proceedings, Boston, MA, AIAA-98-4294.

Garey, M.R., and D.S. Johnson (1979). Computers and Intractability: A Guide to the Theory of NP-Completeness. Bell Laboratories: Murray Hill, New Jersey. pp. 211-214.

Kaber, D.B., M.C. Wright, and M.P. Clamann (2002, March). "Human Response to Adaptive Automation of Early Sensory/Information Acquisition Functions and Later Stages of Information processing in Complex System Operations." Final Report for North Carolina State University, under NASA Langley grant number NAG-1-01039.

Kernstock, N. (1999, October). "A New Era in the Cockpit." *Rotor & Wing*, pp. 16-22.

Kilgard, M.J. (1996, November). "The OpenGL Utility Toolkit Programming Interface, Version 3." Silicon Graphics, Inc.

Kirlik, A., N. Walker, A.D. Fisk, and K. Nagel (1996). "Supporting Perception in the Service of Dynamic Decision Making." *Human Factors*, vol. 38 (2), pp. 288-299.

Latorella, K., and C.L. Jenkins (1999, November 23). "Weather Avoidance Using Route Optimization as a Decision Aid: An AWIN Topical Study." Honeywell Final Report, Agreement No. NCC-1-291.

Layton, C., P.J. Smith, and C.E. McCoy (1994). "Design of a Cooperative Problem-Solving System for En-Route Flight Planning: An Empirical Evaluation." *Human Factors*, vol. 36 (1), pp. 94-119.

Leavitt, C.A. (1996, May 20-23). "Real-Time In-Flight Planning." IEEE National Aerospace and Electronics Conference, Dayton, OH. Vol. 1, A96-37576 10-1. Institute of Electrical and Electronics Engineers: New York. pp. 83-89.

Moore, D.S. (1997). The Active Practice of Statistics. W.H. Freeman and Company: New York. pp. 70-76.

Moray, N., T. Inagaki, and M. Itoh (2000). "Adaptive Automation, Trust, and Self-Confidence in Fault Management of Time-Critical Tasks." *Journal of Experimental Psychology: Applied*, vol. 6 (1). pp. 45-58.

Onken, R. (1997). "The Cockpit Assistant System CASSY as an On-Board Player in the ATM Environment." WEB: http://atm-seminar-97.eurocontrol.fr/onken.html

Parasuraman, R., T.B. Sheridan, and C.D. Wickens (2000, May). "A Model for Types and Levels of Human Interaction with Automation." *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 30 (3), pp. 286-297.

Parasuraman, R. and V.A. Riley (1997). "Humans and Automation: Use, Misuse, Disuse, Abuse." *Human Factors*, vol. 39 (2), pp. 230-253.

Pritchard, D.E. (2000, March). "Dynamic Route Replanning and Retasking of Unmanned Aerial Reconnaissance Vehicles." M.S. Thesis Document, AFIT/GAE/ENY/00M-10. Archived at Air Force Institute of Technology, Wright-Patterson AFB, OH.

Robertson, G. (2000, May 2-4). "Flight Demonstration of an Associate System: A Rotorcraft Pilot's Associate Example." Proceedings from American Helicopter Society Annual Forum.

Sanders, M.S., and E.J. McCormick (1993). Human Factors in Engineering and Design, 7th ed. McGraw-Hill: New York. Chapters 5 and 10.

Sarter, N.B., and D.D. Woods (1995). "How in the World Did We Ever Get into that Mode? Mode Error and Awareness in Supervisory Control." *Human Factors*, vol. 37 (1), pp. 5-19.

Scerbo, M.W. (1996). "Theoretical Perspectives on Adaptive Automation." Automation and Human Performance: Theory and Applications, Eds. R. Parasuraman and M. Mouloua. Lawrence Erlbaum Associates: Mahawah, New Jersey. pp. 37-64.

Schulte, A., P. Stutz, and W. Klockner (1999, September 27-29). "Cognitive Concepts for Tactical Mission Management." Proceeding from the NATO System Concepts and Integration Panel Symposium in Florence, Italy.

Sexton, G.A. (1988). "Cockpit-Crew Systems Design and Integration." Human Factors in Aviation. Eds. E.L. Wiener and D.C. Nagel. Academic Press: New York. pp. 495-526.

StatSoft, Inc. (2002). Electronic Statistics Textbook. Tulsa, OK: StatSoft. WEB: http://www.statsoft.com/textbook/stathome.html

Stringer, J.R. (2002), editor. "Weapons." *Airman*, vol. 46 (1). pp. 44-56.

Taylor, R.M. and J. Reising (1998). "The Human-Electronic Crew: Human-Computer Collaborative Teamworking." NASA no. 19990018356 Conference Paper.

Wickens, C.D., S.E. Gordon, and Y. Liu (1998). Introduction to Human Factors Engineering. Addison-Wesley Educational Publishers: New York. pp. 183-222.

Wiener, E.L. (1988). "Cockpit Automation." Human Factors in Aviation, Eds. E.L. Wiener and D.C. Nagel. Academic Press: New York. pp. 433-461.